



UNIVERSIDADE DE ÉVORA
ESCOLA DE CIÊNCIAS E TECNOLOGIA

Mestrado em Química

Dissertação

**Modelação da Qualidade da Água em
Sistemas de Abastecimento Público**

Susana Rute Guerra Dias

Orientador:

Professor Doutor Henrique Agostinho de Oliveira Moiteiro Vicente

Mestrado em Química

Dissertação

**Modelação da Qualidade da Água em
Sistemas de Abastecimento Público**

Susana Rute Guerra Dias

Orientador:

Professor Doutor Henrique Agostinho de Oliveira Moiteiro Vicente

Ao Rui;

À minha irmã;

Aos meus pais, que o souberam ser.

AGRADECIMENTOS

A execução deste trabalho só foi possível com a colaboração e apoio de várias pessoas, a quem quero prestar o devido agradecimento.

Um agradecimento especial ao Professor Doutor Henrique Agostinho Oliveira Moiteiro Vicente, pelo incansável apoio, disponibilidade e entusiasmo que demonstrou em orientar esta dissertação.

Ao Rui o meu muito obrigado pelo incansável apoio e por toda a ajuda prestada.

À minha família, em especial aos meus pais e irmã, pelo apoio incondicional que me deram.

A todos os meus colegas de trabalho, em especial à Responsável pelo Laboratório, a Dr.^a Maria Leonor Fialho e ao meu colega Gonçalo Candeias, pelo apoio e incentivo durante a elaboração do presente trabalho.

A todos os meus amigos que me foram estimulando pela confiança em mim depositada, fazendo-me acreditar que era possível chegar ao fim com sucesso.

A todos, os meus sinceros agradecimentos.

Modelação da Qualidade da Água em Sistemas de Abastecimento Público

RESUMO

A problemática relacionada com a modelação da qualidade da água destinada ao abastecimento público pode ser abordada de diversos pontos de vista. Neste trabalho recorre-se a metodologias de resolução de problemas que emanam da Área Científica da Inteligência Artificial, assim como a ferramentas utilizadas na procura de soluções como as Árvores de Decisão, as Redes Neurais Artificiais e a Aproximação de Vizinhos.

Actualmente os métodos de avaliação da qualidade da água são muito restritivos já que não permitem aferir a qualidade da água em tempo real. O desenvolvimento de modelos de previsão baseados em técnicas de Descoberta de Conhecimento em Bases de Dados, mostrou ser uma alternativa tendo em vista um comportamento pró-activo que pode contribuir decisivamente para diagnosticar e melhorar a qualidade da água disponibilizada às populações.

No decurso do trabalho, foi utilizada a aprendizagem não-supervisionada tendo em vista encontrar grupos distintos de águas com propriedades físico-químicas semelhantes.

Palavras-chave: Qualidade da Água; Águas de Abastecimento Público; *Data Mining*; Descoberta de Conhecimento em Bases de Dados; Árvores de Decisão; Redes Neurais Artificiais; *k-means*.

Modelling Water Quality in Public Supply Systems

ABSTRACT

The problems related to the modelling of water quality for public supply can be approached from different viewpoints. This work resorts to methods of resolving problems emanating from the Scientific Area of Artificial Intelligence as well as to tools used in the search for solutions such as Decision Trees, Artificial Neural Networks and Nearest-Neighbour Method.

Currently, the methods for assessing water quality are very restrictive because they do not indicate the water quality in real time. The development of forecasting models, based on techniques of Knowledge Discovery in Databases, shows to be an alternative in view of a pro-active behavior that may contribute to diagnose and to improve the quality of water supply to the population.

In this work, unsupervised learning was used to find groups of waters with similar physical and chemical properties.

Keywords: Water Quality; Water Supply; Data Mining; Knowledge Discovery in Databases; Decision Trees; Artificial Neural Networks; k-means.

ÍNDICE

Agradecimentos	v
Resumo	vii
Abstract	ix
Índice	xi
Índice de Figuras	xv
Índice de Tabelas	xvii
Notação	xix
Capítulo I – Introdução.....	1
I.1 – Caracterização da Entidade Acolhedora	4
I.2 – Caracterização das Origens da Água de Abastecimento do Distrito de Évora	7
I.3 – Inovação e Valor Acrescentado.....	9
I.4 – Estrutura e Organização do Trabalho	11
Capítulo II – Análise Química e Microbiológica da Água.....	13
II.1 – Parâmetros Organolépticos	13
II.1.1 – Turvação	13
II.2 – Parâmetros Físico-Químicos.....	14
II.2.1 – pH.....	14
II.2.2 – Condutividade Eléctrica.....	15
II.2.3 – Amónio	16
II.2.4 – Nitrato	16

II.2.5 – Nitrito	17
II.2.6 – Ferro	17
II.2.7 – Oxidabilidade	18
II.2.8 – Sódio	19
II.2.9 – Potássio	19
II.2.10 – Outros Metais	20
II.3 – Parâmetros Microbiológicos	22
II.3.1 – Número de Colónias a 22 °C e a 37 °C	22
II.3.2 – Bactérias Coliformes	23
II.3.3 – <i>Escherichia coli</i>	23
II.3.4 – Enterococos Intestinais	24
II.3.5 – <i>Clostridium perfringens</i>	24
II.4 – Legislação	24
Capítulo III – Descoberta de Conhecimentos em Base de Dados – Modelos e Técnicas	29
III.1 – Descoberta de Conhecimentos em Base de Dados	31
III.2 – <i>Data Mining</i>	35
III.3 – Modelos e Técnicas de <i>Data Mining</i>	38
III.3.1 – Árvores de Decisão	39
III.3.2 – Redes Neurais Artificiais	44
III.3.3 – Aproximação de Vizinhanças	52
III.4 – Avaliação dos Modelos	53
III.4.1 – Avaliação de modelos de classificação	55
III.4.2 – Avaliação de modelos de regressão	58
Capítulo IV – Procedimento Experimental	61
IV.1 – Amostragem e Preservação das Amostras	61
IV.1.1 – Técnicas de Amostragem	63

IV.1.1.1 – Colheita de Amostras de Água para Análise de Níquel, Chumbo e Cobre	63
IV.1.1.2 – Colheita de Amostras de Água para Análises Microbiológicas	63
IV.1.1.3 – Colheita de Amostras de Água para Análises Químicas	64
IV.1.2 – Acondicionamento e Transporte das Amostras	64
IV.1.3 – Pontos e Frequência de Amostragem	64
IV.2 – Métodos e Técnicas Analíticas	66
Capítulo V – Apresentação, Interpretação e Análise de Resultados e Modelos	69
V.1 – Breve descrição da ferramenta WEKA	69
V.2 – Apresentação e Preparação dos Dados	70
V.3 – Modelação	73
V.3.1 – Modelos de Segmentação	75
V.3.2 – Modelos de Regressão	85
V.3.2.1 – Modelos de previsão dos parâmetros nitrato, potássio e sódio	87
V.3.2.2 – Modelos de previsão dos parâmetros nitrato e oxidabilidade	93
V.3.2.3 – Modelos de previsão dos parâmetros nitrato e manganês	102
V.3.2.4 – Modelos de previsão dos parâmetros nitrato e turvação	111
V.3.2.5 – Análise de Sensibilidade dos Modelos Apresentados	119
VI – Conclusões Finais e Perspectivas Futuras	123
Referências Bibliográficas	125
Anexo	I

ÍNDICE DE FIGURAS

Figura 1 – Ciclo biogeoquímico da água.....	2
Figura 2 – Processo de Descoberta de Conhecimento em Bases de Dados	32
Figura 3 – Etapas do processo de DCBD agrupadas em tarefas.	35
Figura 4 – Representação esquemática de neurónio.....	44
Figura 5 – Estrutura geral de um nodo	47
Figura 6 – Arquitecturas de rede	49
Figura 7 – Validação cruzada com dez iterações.....	55
Figura 8 – Distribuição do número de casos pelos concelhos do Distrito de Évora	76
Figura 9 – Associação entre segmentos e concelhos de onde provêm as amostras.....	78
Figura 10 – Árvore de decisão explicativa do modelo de segmentação ($k = 2$).....	78
Figura 11 – Associação entre segmentos e concelhos donde provêm as amostras	80
Figura 12 – Árvore de decisão explicativa do modelo de segmentação ($k = 3$).....	80
Figura 13 – Associação entre segmentos e concelhos donde provêm as amostras	82
Figura 14 – Árvore de decisão explicativa do modelo de segmentação ($k = 4$).....	83
Figura 15 – Articulação entre os segmentos dos diversos modelos de segmentação.....	84
Figura 16 – RNA seleccionada para a previsão de nitrato, potássio e sódio.....	89
Figura 17 – Valores previstos vs valores experimentais das variáveis de saída.....	90
Figura 18 – Resíduos vs valores previstos para as variáveis de saída	91

Figura 19 – Valores previstos vs valores experimentais para os dados de validação.....	93
Figura 20 – Resíduos vs valores previstos relativos ao conjunto de validação	93
Figura 21 – RNAs seleccionadas para previsão de nitrato e oxidabilidade	96
Figura 22 – Valores previstos vs valores experimentais para a topologia 2-19-10-2.....	98
Figura 23 – Valores previstos vs valores experimentais para a topologia 2-20-12-2.....	99
Figura 24 – Resíduos vs valores previstos das topologias 2-19-10-2 e 2-20-12-2	100
Figura 25 – Valores previstos vs valores experimentais para os dados de validação...	101
Figura 26 – Resíduos vs valores previstos relativos ao conjunto de validação	102
Figura 27 – RNAs seleccionadas para a previsão de manganês e nitrato.....	105
Figura 28 – Valores previstos vs valores experimentais para a topologia 2-19-10-2...	106
Figura 29 – Valores previstos vs valores experimentais para a topologia 2-19-17-2...	107
Figura 30 – Resíduos vs valores previstos das topologias 2-19-10-2 e 2-19-17-2.....	109
Figura 31 – Valores previstos vs valores experimentais para os dados de validação...	110
Figura 32 – Resíduos vs valores previstos relativos ao conjunto de validação	110
Figura 33 – RNAs seleccionadas para a previsão de nitrato e turvação.....	113
Figura 34 – Valores previstos vs valores experimentais para a topologia 2-8-6-2.....	115
Figura 35 – Valores previstos vs valores experimentais para a topologia 2-7-4-2.....	116
Figura 36 – Resíduos vs valores previstos das topologias 2-8-6-2 e 2-7-4-2	117
Figura 37 – Valores previstos vs valores experimentais para dados de validação	118
Figura 38 – Resíduos vs valores previstos relativos ao conjunto de validação	119
Figura 39 – Importância relativa das variáveis de entrada	120

ÍNDICE DE TABELAS

Tabela 1 – Valores paramétricos referidos no Decreto-Lei n.º 306/2007	27
Tabela 2 – Etapas do processo de DCBD	33
Tabela 3 – Algumas das definições mais usuais para o termo <i>Data Mining</i>	36
Tabela 4 – Caracterização dos objectivos de <i>Data Mining</i>	37
Tabela 5 – Tarefas e técnicas de <i>Data Mining</i>	39
Tabela 6 – Algumas funções de activação	48
Tabela 7 – Matriz de coincidências 3×3	56
Tabela 8 – Matriz de custos para um problema de classificação com três classes	57
Tabela 9 – Colheita e preservação de amostras de águas	62
Tabela 10 – Parâmetros incluídos no programa de vigilância sanitária	65
Tabela 11 – Parâmetros químicos analisados e métodos de ensaio	66
Tabela 12 – Parâmetros microbiológicos analisados e métodos de ensaio	67
Tabela 13 – Caracterização estatística das variáveis presentes na base de dados	71
Tabela 14 – Número de análises não pedidas ou abaixo do limite de quantificação	72
Tabela 15 – Valores assumidos na implementação de RNAs	74
Tabela 16 – Valores assumidos na indução de árvores de decisão	75
Tabela 17 – Número de objectos por segmento para os vários modelos estudados	76
Tabela 18 – Centros de gravidade dos segmentos obtidos com $k = 2$	77

Tabela 19 – Matriz de coincidências relativa ao modelo apresentado na Figura 10	79
Tabela 20 – Centros de gravidade dos segmentos obtidos com $k = 3$	79
Tabela 21 – Matriz de coincidências relativa ao modelo apresentado na Figura 12	81
Tabela 22 – Centros de gravidade dos segmentos obtidos com $k = 4$	81
Tabela 23 – Matriz de coincidências relativa ao modelo apresentado na Figura 14	84
Tabela 24 – Número de registos com informação completa por par de parâmetros	86
Tabela 25 – Caracterização estatística das variáveis presentes em cada BD	86
Tabela 26 – Valores do DAM, da MQE e de <i>bias</i> para as topologias testadas	88
Tabela 27 – Valores do DAM, da MQE e de <i>bias</i> para as topologias testadas	94
Tabela 28 – Valores do DAM, da MQE e de <i>bias</i> para as topologias testadas	103
Tabela 29 – Valores do DAM, da MQE e de <i>bias</i> para as topologias testadas	111

NOTAÇÃO

AA – Aprendizagem Automática

AD – Árvores de Decisão

BD – Base de Dados

DCBD – Descoberta de Conhecimento em Base de Dados

DM – *Data Mining*

ERSAR – Entidade Reguladora dos Serviços de Águas e dos Resíduos

ETA – Estação de Tratamento de Águas

GC – Gestão do Conhecimento

HPA NHS W5 – Health Protection Agency – National Standard Methods

IA – Inteligência Artificial

IRAR – Instituto Regulador de Águas e Resíduos

ISO – International Standard Organization

nd – Não determinado

NP – Norma Portuguesa

RNAs – Redes Neurais Artificiais

SMEWW – Standard Methods for the Examination of Water and Wastewater

SNIRH – Sistema Nacional de Informação de Recursos Hídricos

VP – Valor paramétrico

WEKA – Waikato Environment for Knowledge Analysis

<Lq – Inferior ao limite de quantificação do método utilizado

CAPÍTULO I – INTRODUÇÃO

A água é indispensável a todas as formas de vida. Sem ela os vegetais e animais não existiriam. As plantas precisam de água para germinar, crescer, florir e dar frutos. O Homem dificilmente sobreviveria uma semana sem beber água. Sendo verdade que as primeiras comunidades da história se organizaram para viver nas proximidades dos cursos de água, também não existem dúvidas que a água continua, nos dias de hoje, a constituir um dos factores mais importantes para o progresso das sociedades contemporâneas. Nenhuma comunidade pode viver ou evoluir sem um abastecimento adequado de água que permita aos seus habitantes viver de modo saudável, confortável e que contribua para o desenvolvimento sustentado da sua economia.

Atendendo ao exposto no parágrafo anterior, é natural que o Homem se preocupe com a forma como a água se distribui na Natureza, constituindo a disponibilidade de água potável um problema crescente da Sociedade actual. Na verdade, para além do abastecimento em quantidade suficiente, é requisito essencial que a água seja inócua do ponto de vista da saúde pública, uma vez que é, também, um dos principais veículos responsáveis pela transmissão/veiculação de doenças.

A água é a única substância que pode surgir na Natureza nos três estados físicos da matéria. A coexistência destes três estados implica que existam transferências contínuas de água entre estados. Este movimento permanente é designado por ciclo da água, ciclo hidrológico ou, mais correctamente, por ciclo biogeoquímico da água. A maior parte da água existente no planeta encontra-se no estado líquido, em mares, rios, lagos, pântanos e na superfície terrestre estando sujeita à evaporação por influência do calor. O vapor de água desta proveniência, juntamente com o proveniente da transpiração dos seres vivos, é libertado para a atmosfera onde sobe até condensar pelo arrefecimento da massa de ar em que está inserido. Ao condensar formam-se pequenas gotículas de água, que vão aumentando de dimensão à medida que mais água condensa ou que várias partículas se agregam, até se tornarem demasiado pesadas e caírem sob a forma de precipitação. No caso do arrefecimento se dar até temperaturas inferiores a 0 °C há a passagem ao estado

sólido, precipitando sob a forma de neve ou sob a forma de granizo dependendo do arrefecimento ser lento ou rápido.

A água que se precipita sobre a superfície terrestre divide-se em duas fracções. Uma parte, cuja importância depende da intensidade da precipitação e das características do solo, penetra neste, enquanto a porção restante se mantém à superfície. Parte da água infiltrada fica retida, tornando-se humidade e, a restante, infiltra-se até encontrar uma camada impermeável, integrando os lençóis freáticos, alimentando nascentes e poços ou descarregando directamente para o mar ou para lagos. A água retida no solo pode perder-se para a atmosfera, por evaporação ou pode ser absorvida pelas raízes de plantas, sendo utilizada no metabolismo e depois devolvida à atmosfera por transpiração.

A água precipitada, que não se infiltra no solo, é incorporada em ribeiros e rios, alimentando lagos, oceanos ou outras formas de armazenamento superficial que, como foi referido anteriormente, está sujeita a evaporação. Fecha-se, deste modo, o ciclo biogeoquímico da água, Figura 1 [Mendes & Oliveira, 2004; Lehr et al., 2005].

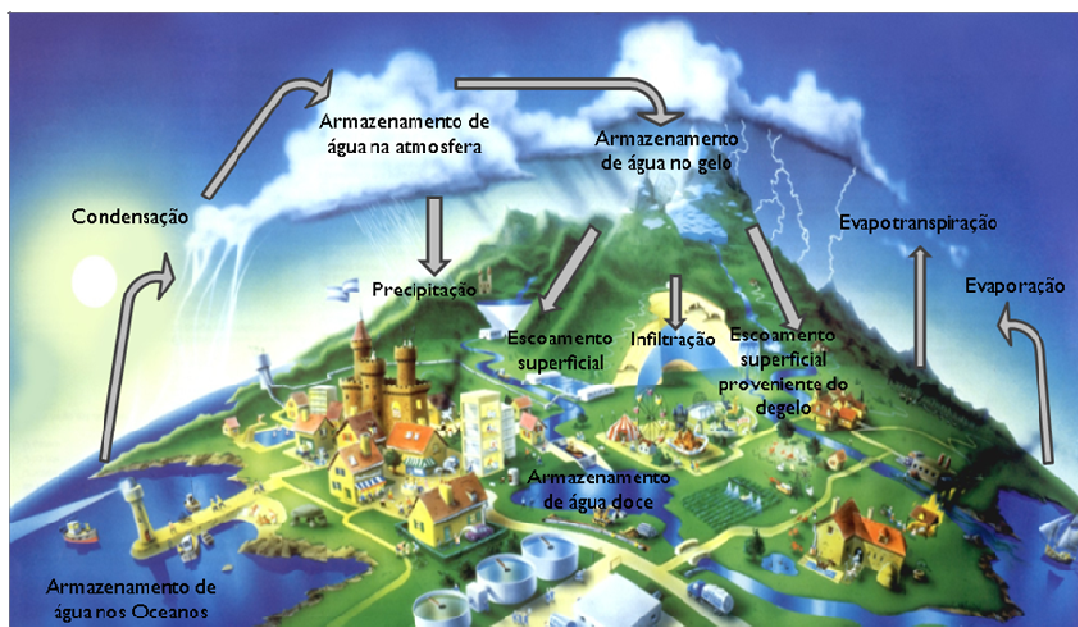


Figura 1 – Ciclo biogeoquímico da água

A problemática da água, ao nível da sua utilização e da gestão dos recursos hídricos disponíveis não é, apenas, um problema quantitativo, enquadrado no seu ciclo biogeoquímico tendo, também, uma faceta qualitativa. A água, enquanto composto dotado de características físico-químicas próprias que afectam e condicionam o seu uso, não se encontra na Natureza no estado puro, i.e. incolor, inodora e insípida. Na verdade, encontra-se associada a outras substâncias que afectam e condicionam os seus possíveis usos. Torna-se, assim, necessário associar à informação da quantidade a indicação da qualidade. A qualidade de uma água pode, então, ser definida como sendo aquilo que caracteriza a sua adaptabilidade para determinados fins, devendo-se ter sempre presente que nenhuma água é de boa qualidade para todos os usos [Spellman, 2008; Varis et al., 2008].

A separação entre a presença habitual e normal de determinadas substâncias, nas águas naturais, e a presença de substâncias não habituais ou em quantidades excessivas, que afectam a sua qualidade e as suas possibilidades de uso, é muito ténue. A diferença é, contudo, substancial. Enquanto a primeira situação é comum e facilmente remediada através quer dos processos de auto-depuração, quer dos processos de tratamento, a segunda situação perturba os ecossistemas sendo, muitas vezes, de difícil remediação inviabilizando, assim, o uso da água [Popek, 2006; Alley, 2007].

O termo “*poluição*” refere-se à inadequação da aplicabilidade da água para o fim desejado. Houve, no entanto, a necessidade de se introduzir o conceito de “*contaminação*”. Deste modo, reserva-se o termo “*poluição*” para referir a presença de substâncias químicas ou de agentes físicos, que provocam prejuízos ou depreciam o valor de uso da água, referindo-se o termo “*contaminação*” à presença de organismos patogénicos ou de substâncias tóxicas que inviabilizem o seu consumo [Popek, 2006; Alley, 2007].

A nível de Saúde Pública existe uma grande preocupação no que diz respeito aos efeitos dos diversos grupos de poluentes e das disfunções ambientais associadas. De entre elas salientam-se [OMS, 2010]:

- Efeitos fisiológicos negativos no Homem e nas espécies cuja sobrevivência a dele dependem;

- Transmissão e libertação de microrganismos patogénicos, responsáveis pela ocorrência de doenças e epidemias diversas;
- Redução da quantidade de oxigénio dissolvido devido quer à biodegradação de substâncias nelas presentes, quer à redução das taxas expectáveis de rearejamento superficial;
- Eutrofização conduzindo a um desenvolvimento excessivo de algas e/ou de plantas superiores imersas, emersas ou flutuantes, acarretando a formação, no fundo, de zonas anaeróbias e a deposição de sedimentos, com formação de cheiros e produção de substâncias tóxicas ou indutoras de sabores inaceitáveis;
- Alteração do funcionamento das cadeias tróficas, pela acumulação de produtos tóxicos em organismos vivos, a qual pode atingir níveis extremamente significativos; e
- Degradação ou destruição de valores estéticos, afectando a qualidade de vida das populações e hipotecando o capital ambiental das gerações futuras.

A maior parte das águas destinadas ao consumo humano têm muitos compostos químicos presentes, sem que tal implique riscos, uma vez que grande parte deles são necessários ou não prejudicam o funcionamento normal do organismo. Na verdade, são os constituintes inorgânicos da água que lhe conferem sabor e valor alimentar. Como exemplo, referimos o cálcio que é um elemento necessário à formação dos dentes, ossos, unhas. Contudo, em quantidades elevadas, pode originar depósitos e incrustações. Outro exemplo característico é o flúor que deverá estar presente numa concentração aproximada de 1 mg/dm^3 para prevenir a cárie dentária mas, quando excede 4 mg/dm^3 , pode originar destruição do esmalte dentário [Mendes & Oliveira, 2004].

I.1 – Caracterização da Entidade Acolhedora

Aos Laboratórios de Saúde Pública (LSPs) compete o apoio analítico às actividades desenvolvidas pelos Serviços de Saúde Pública, tanto ao nível regional como local. O âmbito das suas actividades inclui [ARSA Alentejo, 2010]:

- Participar na elaboração de programas de vigilância e controlo, no âmbito da vigilância sanitária/epidemiológica, da investigação e da cooperação com outras entidades ou sectores;
- Proceder à caracterização físico-química e microbiológica, com as técnicas e métodos que integram os programas de vigilância e controlo, com recurso a meios próprios ou através da contratação de laboratórios de análise devidamente acreditados;
- Proceder ao lançamento dos dados resultantes das referidas caracterizações, tendo em vista a emissão dos relatórios de ensaio;
- Remeter aos serviços competentes os relatórios de ensaio;
- Implementar métodos analíticos, procedimentos de funcionamento e promover a sua actualização;
- Elaborar e manter actualizados os registos dos procedimentos e das análises efectuadas;
- Velar pelo bom estado de conservação, calibração e funcionamento dos equipamentos, instrumentos e utensílios, bem como pela conservação e limpeza das instalações;
- Proceder de forma a garantir a existência, a conformidade e a conservação dos materiais utilizados nos processos de análise da água, bem como elaborar as consultas e proceder à análise de propostas de fornecimento dos referidos materiais;
- Velar e observar o cumprimento das condições de segurança recomendadas; e
- Preparar os recipientes e as embalagens destinadas à recolha de amostras, assim como elaborar os procedimentos adequados.

O grande desafio colocado aos LSPs da Região Alentejo, dos quais fazem parte os LSP's de Évora, Beja e Portalegre, foi adequar a sua missão de forma a responder aos problemas com que se confronta a Saúde Pública e os serviços de Saúde Regionais, com níveis de qualidade que fossem uma referência para a Região.

O LSP de Évora tem como actividade principal a vigilância da qualidade da água dos Sistemas de Abastecimento dos catorze concelhos Distrito de Évora e, ainda, dos concelhos de Alcácer do Sal, Grândola, Santiago do Cacém e Sines do Litoral Alentejano. Executa, ainda, análises de águas de outras proveniências, nomeadamente piscinas recreativas, piscinas de empreendimentos turísticos, piscinas para fins terapêuticos, hemodiálise, balneares, furos e poços. Engloba, neste momento, três sectores distintos (Química de Águas, Microbiologia de Águas e a Micobacteriologia), estando em fase de implementação o sector da Microbiologia de Alimentos.

O Laboratório encontra-se acreditado pelo Instituto Português de Acreditação (IPAC) de acordo com a Norma NP EN ISO/IEC 17025, desde 2006, tendo-lhe sido atribuído o Certificado de Qualidade de Entidade Acreditada com o nº L0424-2, para os ensaios efectuados nos sectores da Química e da Microbiologia de Águas. Faz parte da lista, emitida pelo ERSAR, dos laboratórios aptos para a realização de análises de água destinada ao consumo humano (ensaios microbiológicos e químicos), no âmbito do Decreto-Lei n.º 306/07, de 27 de Agosto.

A acreditação do Laboratório é condição essencial para credibilizar e conferir transparência a todo o processo de avaliação de controlo de qualidade das águas analisadas, para isso o Laboratório possui um quadro de pessoal qualificado, que garante competência no desempenho das suas funções.

O laboratório tem como política a melhoria e actualização contínua dos seus processos para reconhecimento da elevada qualidade dos serviços prestados, baseando-se para isso nas boas práticas profissionais e técnicas implementadas, na perspectiva de atingir o objectivo de satisfazer todos os seus clientes.

Na prossecução desta política a Administração Regional de Saúde do Alentejo, I.P., através do Departamento de Saúde Pública, estabeleceu para o laboratório um Sistema de Gestão baseado na NP EN ISO/IEC 17 025 que tem como objectivos gerais:

- Assegurar a qualidade dos serviços prestados e o permanente exercício de boas práticas profissionais de acordo com os requisitos contratuais, legislação, regulamentos, recomendações e normas aplicáveis à actividade.

- Disponibilizar, manter e actualizar os meios operacionais, logísticos e informáticos necessários ao Sistema de Gestão implementado.
- Envolver todos os colaboradores na manutenção e melhoria contínua do Sistema de Gestão através da sua familiarização com as directrizes e procedimentos que lhe estão associados.

Em complemento aos objectivos acima descritos, são estabelecidos anualmente objectivos quantificáveis.

I.2 – Caracterização das Origens da Água de Abastecimento do Distrito de Évora

As águas para consumo humano, distribuídas no Distrito de Évora, tem origem quer em águas superficiais quer em águas subterrâneas. Neste Distrito encontram-se localizadas as albufeiras de Monte Novo, que abastece os concelhos de Évora e Reguengos de Monsaraz, a albufeira da Vigia, que abastece o concelho de Redondo e a albufeira de Alvito, que abastece os concelhos de Viana do Alentejo e Portel. A composição destas águas é muito variável, dependendo das características do local e da época do ano. De um modo geral, apresentam temperaturas relativamente altas, elevada concentração de matéria orgânica dissolvida, proveniente da decomposição de vegetação e de resíduos de origem antropogénica, elevada turvação devida, principalmente, aos sólidos suspensos (matéria orgânica finamente dividida, microrganismos, plâncton, areias ou argilas), desenvolvimento, por vezes excessivo, de grande variedade de algas, bactérias, cistos e vírus patogénicos que, muitas vezes, contribuem para a existência de sabores e cheiros desagradáveis [Matoso et al., 2004].

Em relação à água de origem subterrânea salienta-se o aquífero Estremoz – Cano, que abastece os concelhos de Alandroal, Borba, Estremoz e Vila Viçosa e o aquífero Viana do Alentejo – Alvito, que abastece os concelhos de Viana do Alentejo e Portel. O aquífero Estremoz – Cano apresenta uma fácies bicarbonatada cálcica a calco-magnesianas. O suporte litológico do sistema aquífero é composto fundamentalmente por dolomitos e mármore, pertencentes ao anticlinal de Estremoz e por calcários ou calcretos, concreções e margas carbonatadas que surgem na região do Cano [Almeida et al., 2000;

Midões & Costa, 2010]. No aquífero Viana do Alentejo – Alvito a fácies dominante é a bicarbonatada cálcica havendo, também, algumas águas com fácies bicarbonatada magnesiana. Trata-se de um sistema aquífero complexo cuja complexidade decorre, sobretudo, da presença de intercalações de rochas menos permeáveis no seio dos calcários [Almeida et al., 2000].

O concelho do Alandroal tem várias zonas de abastecimento de água subterrânea, que provêm do aquífero Estremoz – Cano. Existe uma Estação de Tratamento de Água (ETA) onde a água é misturada e depois tratada.

O concelho de Arraiolos tem nove zonas de abastecimento de água com sistemas de tratamento independentes. Relativamente à origem apenas uma zona provém de uma pequena albufeira tendo, as restantes zonas, origem subterrânea.

O concelho de Borba é abastecido por água subterrânea, que provém do aquífero Estremoz – Cano.

O concelho de Estremoz tem várias zonas de abastecimento de água subterrânea, que provêm do aquífero Estremoz – Cano. Todas elas têm sistemas de tratamento independentes.

No concelho de Évora existem várias zonas de abastecimento que tem origem na albufeira do Monte Novo, sendo o tratamento realizado na ETA existente junto à albufeira. As freguesias de São Brás de Regedouro e de Torre dos Coelheiros constituem as únicas exceções, sendo servidas por água de origem subterrânea proveniente de furos locais.

O concelho de Mora é abastecido por água subterrânea que, do ponto de vista químico, é uma água equilibrada.

O concelho de Mourão é abastecido por água subterrânea proveniente de furos com sistemas de tratamento independentes. É uma zona onde o terreno tem uma grande inclinação e se pratica muita agricultura.

O concelho de Portel é abastecido pela água proveniente da albufeira de Alvito. Tem seis zonas de abastecimento de água superficial. Os fontanários existentes são abastecidos por água subterrânea que tem origem no aquífero Viana do Alentejo – Alvito.

O concelho de Redondo é abastecido por água da albufeira da Vigia, há excepção de Aldeia da Serra que a mistura com água subterrânea proveniente de furos.

O concelho de Reguengos de Monsaraz é abastecido pela albufeira do Monte Novo. No entanto, em períodos de escassez de água ocorre mistura com água subterrânea proveniente de furos.

O concelho de Vendas Novas é abastecido por águas subterrâneas. Tem poucas zonas de abastecimento e, do ponto de vista químico, é considerada uma água equilibrada.

O concelho de Viana do Alentejo tem apenas uma zona de abastecimento servida por sete origens diferentes. Na verdade, antes do tratamento faz-se a mistura de água de origem superficial (albufeira de Alvito) com águas subterrâneas provenientes de vários poços e furos, com origem no aquífero Viana do Alentejo – Alvito.

O concelho de Vila Viçosa tem várias zonas de abastecimento servidas por águas subterrâneas de diversas origens. Parte provém do aquífero Estremoz – Cano e a restante tem origem no furo da Pedreira e noutros furos menos importantes. As diferentes zonas de abastecimento têm sistemas de tratamento independentes.

I.3 – Inovação e Valor Acrescentado

A importância de qualquer tipo de informação para a resolução de um dado problema está directamente relacionada com a frequência com que é considerada e com o modo como é aplicada nessas situações. A fonte de informação ideal seria aquela que facultasse apenas o conhecimento e os dados relevantes para a resolução do problema e exigisse do utilizador apenas um esforço mínimo. As lacunas que as actuais fontes de informação apresentam tornam claro que algo mais é necessário e devem ser o ponto de partida para o desenvolvimento de alternativas.

Pouco se sabe sobre as ferramentas que irão ser utilizadas no futuro. Contudo, algumas das suas características estão, finalmente, a emergir como, por exemplo, o dar resposta a questões complexas em tempo útil. Deste modo, terão que estar ligadas a bases de dados e de conhecimento, terem como base a electrónica e a robótica, serem portáteis e fáceis de utilizar. As ferramentas não poderão ter uma atitude passiva, ou seja, limitarem-se a

responder a questões colocadas pelos utilizadores, mas subscrever um comportamento pró-activo, o que passa por questionar o utilizador, explicar-lhe as suas decisões, procurando ajudar, encontrando uma solução técnica eficiente, desenvolvendo a alternativa mais amiga do ambiente e jamais se assumindo como uma entidade indesejada. São ferramentas com que os decisores irão tratar a informação quer no presente quer no futuro, o que passa por que se estabeleçam ligações entre as diversas fontes de dados e a difusão dos mesmos através da *Internet*, com o decisor a posicionar-se na sua intersecção, não tendo, contudo, que prestar atenção à tecnologia que está a ser utilizada.

Neste trabalho aplicam-se técnicas da Área Científica da Inteligência Artificial ao desenvolvimento de modelos para avaliar a qualidade da água de abastecimento público do Distrito de Évora. Com este trabalho pretende dar-se uma pequena contribuição para o esforço que há a fazer em Portugal, particularmente no Alentejo, na melhoria quer da qualidade da água disponibilizada às populações quer da gestão dos recursos hídricos. Em resumo, os principais objectivos deste trabalho são os seguintes:

- Descrever os problemas e fenómenos associados à qualidade da água de abastecimento público;
- Utilizar as ferramentas da Área Científica da Inteligência Artificial no desenvolvimento de modelos de previsão da qualidade da água de abastecimento público do Distrito de Évora;
- Realçar a importância de tais ferramentas e dos modelos na gestão e exploração sustentada dos recursos hídricos;
- Realçar a capacidade dos modelos na aferição da qualidade da água em tempo real e na descrição de cenários futuros;
- Conhecer quais os parâmetros fundamentais a determinar para efectuar a caracterização completa da qualidade da água; e
- Apresentar modelos para a resolução de problemas relacionados com a avaliação da qualidade da água de abastecimento público.

I.4 – Estrutura e Organização do Trabalho

Esta dissertação está organizada em seis capítulos, não se contabilizando nem as referências bibliográficas nem os anexos. O primeiro é constituído pela presente introdução, que engloba o enquadramento do trabalho, uma breve caracterização do local de trabalho e das origens das águas de abastecimento público do distrito de Évora, apresenta os seus objectivos e descreve a sua estrutura.

No segundo capítulo faz-se referência à legislação aplicada e abordam-se os parâmetros organolépticos, químicos e microbiológicos que, numa perspectiva de um controlo de vigilância, são efectuados pelo Laboratório de Saúde Pública.

No terceiro capítulo é abordada a problemática da Descoberta de Conhecimento em Base de Dados (DCBD), fazendo-se referência aos métodos e técnicas de *Data Mining* utilizados para a avaliação e previsão da qualidade da água de abastecimento público destacando-se, entre outros aspectos, os conceitos, os objectivos e os tipos de abordagens.

Relativamente ao quarto capítulo apresenta-se, numa primeira parte, o procedimento experimental, nomeadamente as metodologias de amostragem e os métodos analíticos utilizados na vigilância da qualidade da água de abastecimento público. Na segunda parte descreve-se a construção da base de dados que irá servir de suporte aos modelos desenvolvidos.

No quinto capítulo fundamentam-se as opções tomadas durante o processo de DCBD e apresentam-se, discutem-se e avaliam-se quer os resultados, quer os modelos construídos.

O sexto e último capítulo apresenta as conclusões finais e perspectiva-se o trabalho que se pretende desenvolver no futuro.

CAPÍTULO II – ANÁLISE QUÍMICA E MICROBIOLÓGICA DA ÁGUA

A água contém, geralmente, diversos componentes que provêm do próprio ambiente natural ou foram introduzidos a partir de actividades humanas. Para caracterizar uma água são determinados diversos parâmetros, que se encontram divididos em parâmetros organolépticos, físico-químicos e microbiológicos.

O Decreto-Lei n.º 306/2007, de 27 de Agosto, atribui à autoridade de saúde a missão de vigilância sanitária das águas dos sistemas de abastecimento público, cabendo-lhe a realização de análises complementares tendo em vista a avaliação do risco para a saúde humana das águas distribuídas às populações. Deste modo, neste capítulo apresenta-se uma breve descrição dos parâmetros cuja análise é efectuada no âmbito da actuação do Laboratório de Saúde Pública de Évora.

II.1 – Parâmetros Organolépticos

Os parâmetros organolépticos (cor, turvação e cheiro) são aqueles cuja apreciação pode ser efectuada pelo técnico de colheita no momento da amostragem, através dos órgãos sensoriais, sendo tomados como parâmetros indicativos. Na verdade, trata-se de parâmetros instáveis que podem sofrer alterações durante o armazenamento e o transporte. No entanto, admite-se que tanto a cor como a turvação não são significativamente alteradas em amostras mantidas a 4 °C e em períodos de armazenamento que não excedam 24 horas. O parâmetro sabor também está incluído nesta categoria. Contudo, só pode ser determinado após ter sido assegurado o controlo bacteriológico da água.

II.1.1 – Turvação

A turvação é uma medida da matéria em suspensão na água. Esta, resulta da presença de substâncias não dissolvidas, finamente dispersas, tais como limo, argilas, grãos de areia,

colóides, plâncton e matéria orgânica [Gray, 2008]. Do ponto de vista da saúde pública, algumas das referidas partículas em suspensão são inócuas. No entanto, podem proteger organismos patogénicos (bactérias, vírus e parasitas) do tratamento de desinfecção [ARS-Norte, 2008].

A turvação pode ocorrer na água das redes públicas devido à libertação de depósitos existentes nas canalizações (provenientes de processos de corrosão), a falhas ocorridas no decurso do tratamento e, ainda, à precipitação do alumínio, utilizado no processo de tratamento, sob a forma de hidróxido de alumínio [Gray, 2008]. Uma água pode ser turva sem que daí resultem riscos para a saúde, contudo, esse facto é interpretado pelo consumidor como um indicador de que a qualidade da água é duvidosa ou de que algo correu mal no processo de tratamento [Mendes & Oliveira, 2004].

II.2 – Parâmetros Físico-Químicos

A análise dos parâmetros físico-químicos visa a identificação dos constituintes de uma água. As impurezas químicas resultam da presença de substâncias dissolvidas que, não raras vezes, condicionam as potenciais utilizações da água. Muitos dos compostos químicos presentes nas águas não implicam riscos para a saúde. Alguns deles são necessários ao organismo. Outros, apesar de não serem essenciais ao funcionamento dos organismos, não causam malefícios desde que presentes em quantidades moderadas. Deste modo, torna-se necessário a quantificação dos parâmetros físico-químicos, de forma a adoptar medidas preventivas sempre que necessário.

II.2.1 – pH

A medida do pH é um dos testes mais importantes e frequentes nas análises de água. O pH de uma água exprime a actividade do ião hidrogénio, indicando a sua acidez ou alcalinidade. O valor de pH de uma água é influenciado pela sua origem e pela natureza dos terrenos que atravessou. As águas naturais apresentam, em geral, valores de pH situados no intervalo entre 6,5 e 8,5 (escala de Sørensen). Em algumas situações, contudo, podem apresentar pH mais elevados como é o caso das águas superficiais, especialmente no Verão, em consequência da actividade biológica das algas. O pH desempenha um papel preponderante na especiação de diversas espécies químicas,

influenciando quer a sua solubilidade, quer a sua toxicidade [Stumm & Morgan, 1996; Huheey et al., 1997; Weiner, 2008].

O pH não tem efeitos directos sobre a saúde do consumidor. Contudo, valores de pH inferiores a sete podem provocar corrosão nas tubagens metálicas que, posteriormente, poderá levar ao aumento do teor em chumbo, cádmio ou outros metais na água para consumo. Valores de pH superiores a oito podem provocar incrustações nas tubagens que, por sua vez, podem desenvolver sabores ou cheiros. Pode, ainda, afectar a eficiência da desinfectação devido à alteração das formas de cloro presentes. O pH pode, ao nível das estações de tratamento de águas, afectar a eficiência dos tratamentos de coagulação, de floculação ou de filtração. O controlo e correcção da acidez são conseguidos adicionando à água adjuvantes alcalinos, como por exemplo, cal ou carbonato de sódio [Mendes & Oliveira, 2004; Gray, 2008].

II.2.2 – Condutividade Eléctrica

A condutividade eléctrica pode ser definida como a capacidade que a água possui para conduzir a corrente eléctrica. Indica a presença de espécies químicas, na forma iónica, permitindo avaliar, de uma forma rápida, a sua mineralização [Clapham, 2004]. A condutividade é muito influenciada pela temperatura, sendo recomendado que as leituras sejam efectuadas a 20 °C.

A condutividade depende da concentração total de iões em solução, da sua natureza, da temperatura e da viscosidade da solução. A origem dos sais é diversa, parte pode resultar de processos de lixiviação dos solos como, por exemplo, os bicarbonatos, carbonatos, cloretos, nitratos, sulfatos e outros sais solúveis de metais como o cálcio, magnésio, potássio e sódio. Outra parte pode provir de efluentes e resíduos agrícolas e/ou industriais que contaminam essas águas. No caso das águas superficiais podem, ainda, provir de sais veiculados por via atmosférica [Mendes & Oliveira, 2004].

A condutividade, dado o seu carácter não específico, não representa um problema para a saúde do consumidor, desde que os limites máximos não sejam atingidos. Contudo, alguns dos componentes podem, em função da sua natureza e características específicas, originar riscos para esse mesmo consumidor. Um outro aspecto que deve ser salientado

está relacionado com o facto de a condutividade condicionar fortemente o uso da água para fins industriais ou de rega [Clapham, 2004; Mendes & Oliveira, 2004].

II.2.3 – Amónio

A presença de amónio no ambiente provém, em geral, de processos de degradação de resíduos de origem vegetal ou animal, sobretudo devidos às actividades agrícolas, pecuárias e industriais. Na água, a contaminação por amónio pode ser um indicador de poluição por adubos, esgotos ou resíduos de animais [ARS-Norte, 2008; Gray, 2008].

A presença de ião amónio na água para consumo humano não tem consequências directas de grande relevância na saúde. No entanto, a sua presença pode comprometer a eficiência da desinfecção uma vez que pode reagir com o cloro residual livre, formando cloroaminas e outros compostos organoclorados, eventualmente cancerígenos. Pode, ainda, promover a formação de nitritos nos sistemas de distribuição, prejudicar o desempenho dos filtros para remoção de manganês, causar problemas ao nível do cheiro e do sabor e provocar corrosão de condutas, em especial as de cobre [Mendes & Oliveira, 2004; ARS-Norte, 2008; Gray, 2008].

II.2.4 – Nitrato

Os nitratos são dos constituintes azotados mais importantes na medida em que, por um lado, são um componente essencial à formação de biomassa das plantas e dos animais e, por outro lado, são um poluente importante das águas brutas utilizadas na produção de água para consumo humano.

A quantidade de nitrato presente nas águas aumenta quer com a quantidade de fertilizantes azotados aplicados no solo quer com a descarga de efluentes/resíduos industriais, agrícolas e domésticos. Em regiões onde o solo é contaminado com águas residuais domésticas ou provenientes da criação animal, a conversão biológica pode levar à formação de nitratos que se infiltram em profundidade e se acumulam nos lençóis freáticos. A estes também se vão juntar alguns contributos de origem industrial directa (no fabrico de explosivos, como oxidante na indústria química, ou como conservantes na indústria alimentar) [Gray, 2008].

A presença de teores elevados nitratos na água para consumo humano pode originar problemas sanitários graves, nomeadamente no caso dos bebés, devido à formação de metahemoglobina, responsável pelo “síndrome do bebé azul”. O nitrato é reduzido a nitrito no estômago da criança, onde o suco gástrico é menos ácido que o dos adultos, que, por sua vez, vai oxidar a hemoglobina a metahemoglobina inviabilizando o transporte do oxigénio, resultando na morte do bebé por asfixia, acompanhada de uma coloração azul característica (cianose) [ARS-Norte, 2008; Gray, 2008].

No que respeita ao ambiente em geral e à qualidade das águas em particular, a presença de nitratos (juntamente com outros nutrientes, nomeadamente fosfatos) é responsável pela eutrofização das massas de água [Mendes & Oliveira, 2004].

II.2.5 – Nitrito

Os nitritos existentes numa água resultam, de um modo geral, da oxidação do azoto amoniacal ou da redução dos nitratos. Em condições normais, a conversão de nitritos em nitratos é quase imediata. Deste modo, a sua presença reflecte a existência de processos inibitórios da formação de nitratos [Gray, 2008].

O aparecimento de nitritos nas águas pode ter origem biológica, resultante da redução microbiana dos nitratos ou química, por oxidação do amoníaco proveniente da esterilização/desinfecção das águas por cloroaminas, em especial a temperaturas elevadas [Mendes & Oliveira, 2004; ARS-Norte, 2008; Gray, 2008].

Ao nível da saúde pública, como foi referido anteriormente, inibem o transporte do oxigénio, podendo ocorrer morte por asfixia. Outros riscos associados à presença de nitritos e nitratos nas águas estão relacionados com a possibilidade destes reagirem com aminas e amidas provenientes da hidrólise de proteínas formando, no estômago, nitrosaminas e nitrosamidas que, em muitos casos, possuem propriedades cancerígenas [Mendes & Oliveira, 2004; Gray, 2008].

II.2.6 – Ferro

O ferro é um elemento muito abundante na crosta terrestre. O ferro pode ocorrer em solução (livre ou na forma de complexos) ou nos sedimentos. Em condições redutoras,

predomina na forma de ferro (II). Quando há exposição ao ar ou na presença de oxidantes passa à forma de ferro (III) podendo ser hidrolisado e precipitado. Contudo, na presença de complexantes ou a pH baixos o ferro pode manter-se dissolvido. A concentração de ferro nas águas superficiais é normalmente baixa. A presença de ferro na água pode resultar da lixiviação dos solos, de descargas industriais, actividade mineira, corrosão e actividade microbiana [Gray, 2008].

Muito embora o ferro não apresente inconvenientes à saúde nas concentrações normalmente encontradas nas águas naturais, pode provocar problemas de ordem estética (manchas em roupas ou em loiças), prejudicar determinados usos industriais ou conferir sabor desagradável à água [Mendes & Oliveira, 2004; Gray, 2008].

II.2.7 – Oxidabilidade

A oxidabilidade permite avaliar a quantidade de matéria orgânica presente na água que é quimicamente oxidável. É proveniente da acumulação de substâncias orgânicas de origem biológica (constituídas fundamentalmente por proteínas, gorduras e açúcares) e/ou originadas por actividades antropogénicas. A oxidabilidade constitui, assim, um critério de controlo de poluição, sendo um indício de contaminação química e bacteriológica. Na verdade, o conjunto de substâncias oxidáveis pelo permanganato de potássio permite avançar com uma estimativa relativa à quantidade de matéria orgânica presente na água.

Algumas substâncias orgânicas podem estar, directamente, na origem de formação de cores e sabores desagradáveis ou, indirectamente, por permitirem o desenvolvimento de organismos (algas, cianobactérias, fungos, bactérias e actinomicetas) responsáveis por secreções, que as originam [ARS-Norte, 2008].

A presença de substâncias orgânicas oxidáveis, numa água destinada ao consumo humano, pode não representar um risco específico em termos sanitários. Contudo, o aumento da oxidabilidade poderá ser indicador da presença de poluentes que podem apresentar riscos para o consumidor. Um exemplo típico é a formação de trihalometanos em águas para consumo humano que resultam reacção entre a matéria orgânica e o cloro [Gray, 2008].

II.2.8 – Sódio

O sódio é um elemento alcalino muito abundante na crosta terrestre. Devido à sua elevada reactividade encontra-se sempre sob a forma combinada. Trata-se de um elemento essencial para muitos organismos que, no entanto, pode ser tóxico para muitas plantas e animais quando presente em concentrações elevadas [Huheey et al., 1997].

A concentração de sódio nas águas subterrâneas pode variar com a profundidade do poço e alcançar maiores concentrações do que nas águas superficiais. Pode ser proveniente de esgotos domésticos, efluentes industriais ou fontes naturais de sais de sódio. Nas regiões junto ao litoral, os lençóis freáticos podem sofrer intrusões marinhas, responsáveis por níveis de sódio elevados [Gray, 2008].

Teores excessivos de sódio podem levar a situações de risco. Nos bebés, o sódio provoca problemas devido à falta de capacidade dos rins para a recepção de doses elevadas deste composto. Pode, ainda, originar problemas gastrointestinais com diarreias associadas, que provocam desidratações. No caso dos adultos teores de sódio elevados podem causar problemas cardiovasculares ligados à hipertensão, trombozes da coronária, cardiopatias ligadas à arteriosclerose, entre outras. Os doentes renais devem seguir um regime hipossódico. Relativamente às grávidas, a ingestão de sódio pode provocar retenções de líquidos, em especial a nível dos membros inferiores [ARS-Norte, 2008; Gray, 2008].

II.2.9 – Potássio

O potássio é um metal alcalino muito abundante na natureza sendo, normalmente, encontrado na forma iónica. Este metal surge naturalmente nas águas, variando o seu teor com os contextos geológicos envolventes. Pode, também, ocorrer devido a contaminações provenientes de descargas originadas, por exemplo, pela indústria extractiva, fabrico de fertilizantes e fabrico de vidro [Mendes & Oliveira, 2004].

O potássio é um elemento essencial à vida intervindo ao nível do equilíbrio osmótico da célula, da transmissão de influxos nervosos e em sistemas enzimáticos ligados ao metabolismo dos glúcidos. A deficiência em potássio provoca sintomas de fraqueza muscular e perda da capacidade cerebral [Huheey et al., 1997].

II.2.10 – Outros Metais

O manganês é um dos metais presentes na crosta terrestre ocorrendo, normalmente, associado ao ferro. Este elemento é essencial ao Homem e outros animais estando presente em variados alimentos. Encontra-se, muitas vezes, nas águas particularmente em condições anaeróbias ou de baixa oxidação [ARS-Norte, 2008].

O manganês é um elemento químico essencial à vida, necessário em pequenas quantidades, intervindo no funcionamento do metabolismo humano, nomeadamente no crescimento e no metabolismo dos glúcidos e dos lípidos. Os efeitos adversos provenientes do manganês podem resultar tanto devido a deficiência ou excesso deste elemento. Causa danos ao nível neurológico após exposição por via respiratória ou após ingestão de água com elevados níveis de contaminação [Huheey et al., 1997; Gray, 2008].

O cádmio é um metal pouco frequente na crosta terrestre encontrando-se, na maior parte dos solos, em quantidades vestigiais. A sua presença no ambiente tem origem industrial ligada, sobretudo, às actividades extractivas. Juntamente com o níquel é muito utilizado em pilhas recarregáveis [Mendes & Oliveira, 2004].

O cádmio não é um elemento essencial ao Homem sendo tóxico, mesmo em pequenas quantidades. Provoca atrasos no crescimento, hipertensão e acidentes cardiovasculares, acumulando-se nos rins, no fígado ou órgãos reprodutivos. O cádmio substitui o zinco e liga-se, irreversivelmente, a proteínas originando carcinomas e dermatoses [Huheey et al., 1997; Gray, 2008].

O crómio é um elemento pesado, litofílico, com características fortemente ácidas. Para alguns organismos o crómio é essencial encontrando-se presente em quantidades diminutas nas águas. Em geral, está presente nas reservas de água na forma de crómio (III) e de crómio (IV). Pode surgir como resultado de descargas industriais uma vez que é muito utilizado em processos industriais, sendo utilizado, entre outros, no fabrico de aços especiais, tratamento de superfícies e no fabrico de ligas. As águas tratadas apresentam, em geral, valores reduzidos, quando isso não se verifica é devido a processos de corrosão do crómio das tubagens [Mendes & Oliveira, 2004].

O crómio é cancerígeno para o Homem quando absorvido pelas vias respiratórias, em solução a sua toxicidade depende do seu estado de valência, sendo o crómio hexavalente mais tóxico que o crómio trivalente [Huheey et al., 1997; Gray, 2008].

O cobre é um elemento abundante na crosta terrestre. Na forma metálica é muito utilizada em ligas e na indústria. As canalizações em cobre são muito utilizadas nas redes de distribuição de água devido à sua resistência à corrosão. Na agricultura, utilizam-se sais de cobre incorporados em insecticidas, fungicidas e algicidas. A contaminação das águas naturais com cobre resulta das actividades humanas pois embora o metal seja muito pouco solúvel, muitos dos seus sais são bastante solúveis [Mendes & Oliveira, 2004].

O cobre é um elemento essencial à vida enquanto activador de sistemas enzimáticos ligados, sobretudo, à cadeia respiratória. Contribui, tal como o ferro, para a regeneração da hemoglobina. Em excesso pode provocar inflamações gastrointestinais crónicas. Embora o teor de cobre seja, normalmente, reduzido e não implique riscos para a saúde pública verifica-se que a presença de concentrações superiores a 1 mg/dm^3 pode provocar um sabor metálico e a turvação da água, coloração das loiças sanitárias e enegrecimento de alguns alimentos no decurso da cozedura [Huheey et al., 1997; Gray, 2008].

O níquel é um metal pesado, calcófilo e siderófilo com características básicas, muito utilizado num grande número de indústrias, desde fabrico de ligas, cobertura de superfícies, indústria cerâmica e fabrico de vidros. Também é muito utilizado no fabrico de objectos decorativos e em revestimentos para protecção contra a corrosão. A sua presença nas águas é, sobretudo, de origem antropogénica apresentando, geralmente, baixos teores. Nas águas tratadas, os teores de níquel podem ser mais elevados, devido ao contacto desta com elementos niquelados utilizados nas tubagens [Mendes & Oliveira, 2004].

O níquel é relativamente pouco tóxico por via digestiva devido à sua baixa assimilabilidade intestinal. É um elemento essencial participando na actividade enzimática, na acção hormonal, assim como no metabolismo em geral. Pode provocar dermatites por contacto directo e, por inalação, cancro nos pulmões e outros problemas respiratórios [Huheey et al., 1997; Gray, 2008].

O chumbo é um metal pesado, calcófilo e de características anfotéricas. O chumbo apresenta-se na água no estado de oxidação +2 e surge como resultado de descargas industriais, das actividades extractivas, restos de combustíveis e a partir da dissolução das canalizações mais antigas. Algumas tubagens de plástico estabilizado com chumbo constituem, também, uma fonte de contaminação. Quando em solução, o chumbo é incolor, inodoro e insípido o que faz com que, mesmo em elevadas concentrações, não seja perceptível [Mendes & Oliveira, 2004].

O chumbo é um metal extremamente tóxico e tem efeitos cumulativos. Provoca disfunções ao nível dos rins, aparelho reprodutor, fígado, cérebro, sistema nervoso central e pode conduzir à morte. Faz parte da tríade de poluentes da água, a que normalmente se associam os maiores riscos para a saúde pública, juntamente com o mercúrio e o cádmio [Huheey et al., 1997; Gray, 2008].

O alumínio é o terceiro elemento mais abundante na superfície terrestre. Contudo, a sua solubilidade faz com que raramente seja uma preocupação para os sistemas de abastecimento de água. A principal fonte de alumínio nas águas de consumo humano advém do uso de sulfato de alumínio como coagulante nas estações de tratamento. Contudo, o alumínio é, de um modo geral, excretado. [Huheey et al., 1997; ARS-Norte, 2008; Gray, 2008].

II.3 – Parâmetros Microbiológicos

Na água podem-se encontrar numerosos e diversos microrganismos. Alguns não representam qualquer perigo para o Homem enquanto outros, pelo contrário, podem ser causadores ou transmissores de doenças. Deste modo, torna-se necessário adoptar medidas preventivas que possam garantir que a sua presença seja minimizada na água de abastecimento público como, por exemplo, o recurso ao cloro e seus derivados, ao ozono ou a radiações ultravioletas [ARS-Norte, 2008; Gray, 2008].

II.3.1 – Número de Colónias a 22 °C e a 37 °C

A contagem do número de colónias a 22 °C e a 37 °C é uma análise que se faz frequentemente tendo em vista o controlo microbiológico de águas para consumo

humano. A contagem do número total de germes tem pouco significado como indicador de poluição fecal. Contudo, quando efectuada regularmente, permite controlar a variação quantitativa dos organismos presentes, servindo a temperatura como factor selectivo entre os organismos mesófilos (22 °C) e aqueles que apresentam uma maior resistência a temperaturas mais elevadas (37 °C). Qualquer aumento súbito do número de colónias pode constituir um primeiro aviso de poluição e levar a medidas imediatas [ARS-Norte, 2008].

Após o tratamento, a presença destes microrganismos indica a existência de problemas, devendo agir-se de forma preventiva, assegurando uma desinfecção eficaz. Estes organismos não causam doenças em pessoas saudáveis podendo, no entanto, contribuir para introduzir alterações da cor, cheiro e sabor da água [Gray, 2008].

II.3.2 – Bactérias Coliformes

A presença de bactérias coliformes numa água para consumo humano torna essa água inaceitável. Constitui uma indicação de que poderão estar presentes bactérias patogénicas com origem fecal. A sua presença indica que o tratamento foi inadequado e/ou que se verificou uma contaminação posterior, ao nível da distribuição. As bactérias coliformes surgem tanto nos esgotos como nas águas naturais sendo capazes de se multiplicarem [ARS-Norte, 2008; Gray, 2008].

As doenças causadas por este tipo de bactérias podem incluir a febre tifóide, gastroenterites víricas e bacterianas e ainda hepatite A [Gray, 2008].

II.3.3 – *Escherichia coli*

A *Escherichia coli* é, sem dúvida, um organismo de origem fecal encontrando-se, tal como os outros coliformes, espalhados pela natureza. Está presente no tracto intestinal e fezes dos seres humanos e dos animais de sangue quente. A sua presença na água para consumo humano juntamente com níveis elevados de nitrato e/ou cloreto indica, normalmente, contaminação por águas residuais. A *Escherichia coli* produz uma toxina nociva, podendo causar danos graves. A infecção provoca diarreia, normalmente com presença de sangue, e dores abdominais [ARS-Norte, 2008; Gray, 2008].

II.3.4 – Enterococos Intestinais

Os Enterococos são outro indicador de contaminação fecal. São residentes habituais do tracto intestinal do Homem e de muitos outros animais podendo ser utilizados como indicadores da eficiência do tratamento. Os Enterococos intestinais englobam um vasto número de espécies tais como os *Enterococos faecalis*, *E. faecium*, *E. durans*, *E. hirae*.

As doenças causadas pelos Enterococos incluem febre tifóide, gastroenterites víricas e bacterianas e a hepatite A [ARS-Norte, 2008; Gray, 2008].

II.3.5 – *Clostridium perfringens*

O *Clostridium perfringens* tem um papel auxiliar na determinação da qualidade da água. Encontra-se, habitualmente, no tracto intestinal dos animais e dos humanos, formando esporos que são mais resistentes ao aquecimento quando comparados com as células vegetativas. O facto de formarem esporos traduz-se numa vantagem para a sua detecção na água, uma vez que estes têm a capacidade de sobreviver durante muito mais tempo, indicando poluição fecal remota ou intermitente [ARS-Norte, 2008; Gray, 2008].

Os esporos nem sempre são inactivados pela desinfecção com cloro. Contudo, não constituem um perigo para a saúde pública. No entanto, se encontram condições favoráveis ao crescimento, adquirem a forma vegetativa e então tornam-se perigosos para a saúde. A transmissão para o Homem faz-se através de alimentos contaminados com águas onde a bactéria seja encontrada. Produzem toxinas que dão lugar ao aparecimento de gastroenterites [ARS-Norte, 2008; Gray, 2008].

II.4 – Legislação

A 5 de Setembro de 2001 foi publicado o Decreto-Lei n.º 243/2001 que transpôs, para ordem jurídica interna, a Directiva n.º 98/83/CE do Conselho da Europa, de 3 de Novembro, relativa à qualidade da água destinada ao consumo humano. Manteve os aspectos fundamentais do anterior diploma, o Decreto-Lei n.º 236/98, de 1 de Agosto, que já definia o essencial das obrigações das entidades gestoras, nomeadamente a apresentação do programa de controlo da qualidade da água para consumo humano, a

frequência de amostragem de acordo com a população servida, a comunicação dos incumprimentos de valores paramétricos e de outras situações que comportassem risco para a saúde humana, a publicação dos dados analíticos da implementação do programa de controlo da qualidade da água relativos ao ano transacto, a realização de análises preferencialmente em laboratórios de ensaios credenciados e os métodos analíticos de referência.

Relativamente ao anterior diploma legal, o Decreto-Lei n.º 243/2001, de 5 de Setembro, modificou a lista dos parâmetros a realizar, alterou alguns valores paramétricos, abordou de uma forma mais racionalizada o controlo de pesticidas, estabeleceu que o controlo da qualidade da água passava a ser feito na torneira do consumidor e definiu a necessidade de regulamentação das situações em que a gestão e a exploração de um sistema de abastecimento público de água estão sob a responsabilidade de duas ou mais entidades gestoras.

A alteração mais significativa foi, sem dúvida, a criação de uma autoridade responsável pela coordenação da implementação e da fiscalização da aplicação do referido diploma, o Instituto Regulador de Águas e Resíduos (IRAR), actualmente designado por Entidade Reguladora dos Serviços de Água e dos Resíduos (ERSAR, IP). Procedeu-se, assim, à concentração de um conjunto essencial de atribuições anteriormente dispersas por várias entidades públicas, que dificultava uma maior eficiência da Administração na fiscalização de uma matéria essencial à protecção da saúde humana.

A necessidade de, por um lado, adaptar melhor a legislação nacional relativa à qualidade da água para consumo humano à Directiva n.º 98/83/CE, do Conselho da Europa, de 3 de Novembro e, por outro lado, a necessidade de se fazer um balanço rigoroso sobre a implementação Decreto-Lei n.º 243/2001, de 5 de Setembro esteve na base da sua revisão.

O Decreto-Lei n.º 306/2007, de 27 de Agosto, define uma abordagem mais racionalizada para as zonas de abastecimento com volumes médios diários inferiores a 100 m³, nomeadamente no que diz respeito à frequência de amostragem. Introduziu novos parâmetros no controlo da qualidade da água, tendo em conta a existência, em algumas zonas do País, de águas com dureza elevada ou agressivas, ou com frequente aparecimento de fluorescências de cianobactérias, razões pelas quais deverão ser

controladas através da análise de parâmetros específicos e, por fim, tem em conta que a água para consumo humano pode ser fornecida através de sistemas públicos ou particulares de abastecimento, regulamentando estes últimos.

Com base no Decreto-Lei n.º 306/2007, de 27 de Agosto, entende-se por “*água destinada ao consumo humano*” *toda a água no seu estado original, ou após tratamento, destinada a ser bebida, a cozinhar, à preparação de alimentos, à higiene pessoal ou a outros fins domésticos, independentemente da sua origem e de ser fornecida a partir de uma rede de distribuição, de um camião ou navio-cisterna, em garrafas ou outros recipientes, com ou sem fins comerciais, e ainda toda a água utilizada numa empresa da indústria alimentar para fabrico, transformação, conservação ou comercialização de produtos ou substâncias destinados ao consumo humano, assim como a utilizada na limpeza de superfícies, objectos e materiais que podem estar em contacto com os alimentos, excepto quando a utilização dessa água não afecta a salubridade do género alimentício na sua forma acabada.*

O Decreto-Lei n.º 306/2007, de 27 de Agosto, estabelece no n.º 1 do artigo 6.º que a água para consumo humano deve respeitar os valores paramétricos que constam das partes I, II e III do anexo I. Na parte I constam os valores paramétricos referentes aos parâmetros microbiológicos, a parte II refere os valores paramétricos relativos aos parâmetros químicos e, por fim, na parte III constam os parâmetros indicadores que, por definição, são os parâmetros cujo valor deve ser considerado como valor guia.

Os valores paramétricos estabelecidos tendo em vista o controlo de qualidade da água para consumo humano encontram-se na Tabela 1. Salienta-se, no entanto, que na Tabela 1 não se apresentam todos os parâmetros mencionados na referida legislação mas, apenas, os parâmetros determinados no Laboratório de Saúde Pública de Évora. O termo ‘sem alteração anormal’ significa, com base num histórico de análises, resultados dentro dos critérios estabelecidos pelas entidades gestoras. No entanto, o referido Decreto-Lei recomenda que o número de colónias a 22 °C e a 37 °C não ultrapassem, respectivamente, cem e vinte.

Tabela 1 – Valores paramétricos dos parâmetros microbiológicos, químicos e indicadores de acordo com o Decreto-Lei n.º 306/2007, de 27 de Agosto

Parâmetros		Valor Paramétrico	Unidade
Microbiológicos	<i>Escherichia coli</i>	0	Número / 100 cm ³
	Enterococos	0	Número / 100 cm ³
Químicos	Cádmio	5,0	µg dm ⁻³
	Crómio	50	µg dm ⁻³
	Cobre	2,0	µg dm ⁻³
	Chumbo	25	µg dm ⁻³
	Níquel	20	µg dm ⁻³
	Nitrato	50	mg dm ⁻³
	Nitrito	0,5	mg dm ⁻³
Indicadores	Alumínio	200	µg dm ⁻³
	Amónio	0,50	mg dm ⁻³
	<i>Clostridium perfringens</i>	0	Número / 100 cm ³
	Condutividade	2500	µS cm ⁻¹ a 20°C
	pH	6,5 ≤ pH ≤ 9,0	Escala Sørensen
	Ferro	200	µg dm ⁻³
	Manganês	50	mg dm ⁻³
	Oxidabilidade	5	mg dm ⁻³
	Sódio	200	mg dm ⁻³
	Número de colónias 22°C	sem alteração anormal	Número / cm ³
	Número de colónias 37°C	sem alteração anormal	Número / cm ³
	Bactérias coliformes	0	Número / 100 cm ³
	Turvação	4	UNT

CAPÍTULO III – DESCOBERTA DE CONHECIMENTOS EM BASE DE DADOS – MODELOS E TÉCNICAS

A palavra informática, introduzida pela primeira vez em 1962 por Philippe Dreyfus, resulta da combinação das palavras **informação** e **automática**, tendo sido definida como uma técnica de tratamento automático e racional da informação considerando-se, neste contexto, a informação como o suporte do conhecimento e da comunicação nos domínios técnico, económico e social. A academia francesa concede-lhe, em 1967, o estatuto de Ciência definindo, em simultâneo, computador como máquina automática capaz de realizar operações matemáticas e lógicas com aplicações científicas, administrativas ou contabilísticas [Birrien, 2008].

A interacção entre o computador, suporte da informação, os programas capazes de tratar a informação e a componente humana das organizações que se relaciona directamente com os programas, está na base do conceito de Tecnologias da Informação (TI), também designadas por Tecnologias da Informação e Comunicação. As TI rapidamente se afirmaram, promovendo a melhoria do tratamento da informação, quer no plano da recolha/armazenamento, quer no plano da selecção/distribuição.

As organizações, operando em ambientes cada vez mais competitivos à escala global, rapidamente reconheceram as potencialidades das TI e depressa as integraram tendo em vista a implementação de novas estratégias, a melhoria de processos, a introdução de novos produtos e a satisfação dos clientes. Na verdade, a eficácia patenteada pelas TI no que concerne à recolha e armazenamento de dados fez com que nas organizações de todo o mundo nascesse a ideia que uma maior quantidade de dados gera vantagens competitivas. No entanto, o facto dos dados, na maioria dos casos, não serem tratados em tempo útil ou a utilização de técnicas inadequadas leva a que o resultado obtido nem sempre seja o desejado. Torna-se, pois, vital analisar os dados, extrair conhecimento, generalizar, sublinhar princípios e estruturar a informação. Não basta ter a informação disponível. A obtenção de mais-valias está fortemente correlacionada com a capacidade

de seleccionar a informação certa para o problema específico [Perlovsky, 2001; Turban et al., 2004; Soman et al., 2006].

A análise de dados não é uma discussão recente. Na verdade, desde há muito que ficou claro que o cérebro humano analisa dados, trata informação e adquire conhecimento de modo diferente, através de processos de aprendizagem. Numa perspectiva sistémica a aprendizagem pode ser definida como as alterações do sistema que lhe permitem refazer as mesmas tarefas de uma forma mais eficaz e eficiente no futuro [Adriaans & Zantinge, 1996; Hand et al., 2001; Russell & Norvig, 2003; Costa & Simões, 2008]. A aprendizagem pode envolver duas fases distintas de processamento da informação – a fase indutiva e a fase dedutiva. Na primeira tem-se a aprendizagem baseada em casos, na qual os padrões e as regras são determinados a partir dos dados e das experiências. Na segunda forma, as regras e os padrões são utilizados para criar novos factos.

No que concerne à aprendizagem deve ser realçada a relação entre o modelo de aprendizagem e o seu ambiente. Neste âmbito existem, fundamentalmente, três paradigmas de aprendizagem [Patterson, 1996]:

- Aprendizagem supervisionada – a aprendizagem é levada a cabo a partir de um conjunto de casos de treino ou exemplos onde, cada um deles, é formado por um vector de entrada e um vector de saída ou de resposta. Deste modo, são fornecidas ao sistema as respostas correctas;
- Aprendizagem de reforço – nesta técnica não se apresenta a resposta correcta ao sistema. Apenas se dá uma indicação sobre se a resposta apresentada pelo sistema é correcta ou não, sendo esta informação utilizada no melhoramento do seu desempenho; e
- Aprendizagem não supervisionada – neste caso não chega ao sistema qualquer indicação externa sobre a resposta correcta. A aprendizagem é feita a partir de descobertas de características nos dados de entrada, procurando regularidades estatísticas ou agrupamentos de padrões.

III.1 – Descoberta de Conhecimentos em Base de Dados

As Tecnologias de Informação permitiram, como foi referido anteriormente, o crescimento exponencial da quantidade de dados armazenados quer em termos numéricos, quer em termos de complexidade. Como consequência deste aumento efectivo de informação, o seu processamento através dos métodos tradicionais tornou-se cada vez mais difícil e complexo [Michalski et al., 1998; Goebel & Gruenwald, 1999; Sousa, 2004]. Deste modo, começaram a surgir aplicações vocacionadas para a tarefa de Descoberta de Conhecimento em Bases de Dados incorporando ferramentas de *Data Mining* para extrair informação, padrões de comportamento e tendências desconhecidas de forma autónoma e porventura automática [Thuraisingham, 1999; Apte et al., 2002; Cios et al., 2007; Taniar, 2008]. Esta é uma área, em progressiva expansão que visa estudar, investigar e desenvolver processos que possibilitem extrair conhecimento inovador, útil e válido implicitamente incluído em grandes volumes de dados e facultá-lo ao utilizador numa forma acessível e clara [Sousa, 2004].

O termo Descoberta de Conhecimento em Bases de Dados (DCBD), formalizado em 1989, traduz o conceito de procura de conhecimento em Bases de Dados (BD) e refere-se a um processo que envolve a identificação e o reconhecimento de padrões de uma forma automática, i.e., a obtenção, sem nenhuma formulação prévia de hipóteses, de informações relevantes e desconhecidas que podem ser úteis no processo de tomada de decisão [Fayyad et al., 1996a; Sousa, 2004].

O processo de Descoberta de Conhecimento em Bases de Dados depende de uma nova geração de ferramentas e técnicas de análise de dados, envolvendo diversas etapas. Consoante o autor, a área de especialização e a abordagem assim se consideram mais ou menos etapas. Fayyad e seus colaboradores [Fayyad et al., 1996b] distinguem nove etapas, nomeadamente a definição do domínio de aplicação, dos conhecimentos relevantes e dos objectivos que se pretendem alcançar; a criação/selecção de um conjunto de dados alvo; a limpeza e pré-processamento dos dados; a redução e projecção dos dados; a escolha do método de *Data Mining* (DM) adequado; a selecção da técnica (algoritmo)

de *Data Mining* apropriada; a etapa *Data Mining*; a interpretação dos padrões obtidos e, por fim, a utilização do conhecimento obtido.

Adriaans e Zantinge [Adriaans & Zantinge, 1996], por seu lado, referem apenas cinco etapas, especificamente a definição do problema; o enriquecimento dos dados; a codificação dos dados; a etapa de *Data Mining* e, por último, a sistematização e divulgação do novo conhecimento.

O exposto nos parágrafos anteriores mostra que o processo de Descoberta de Conhecimento em Bases de Dados apresenta etapas comuns, designadamente a selecção, o pré-processamento dos dados, a transformação dos dados, *Data Mining* e interpretação. Na Figura 2 ilustra, de uma forma esquemática, as etapas do processo de Descoberta de Conhecimento em Bases de Dados que se caracterizam na Tabela 2.

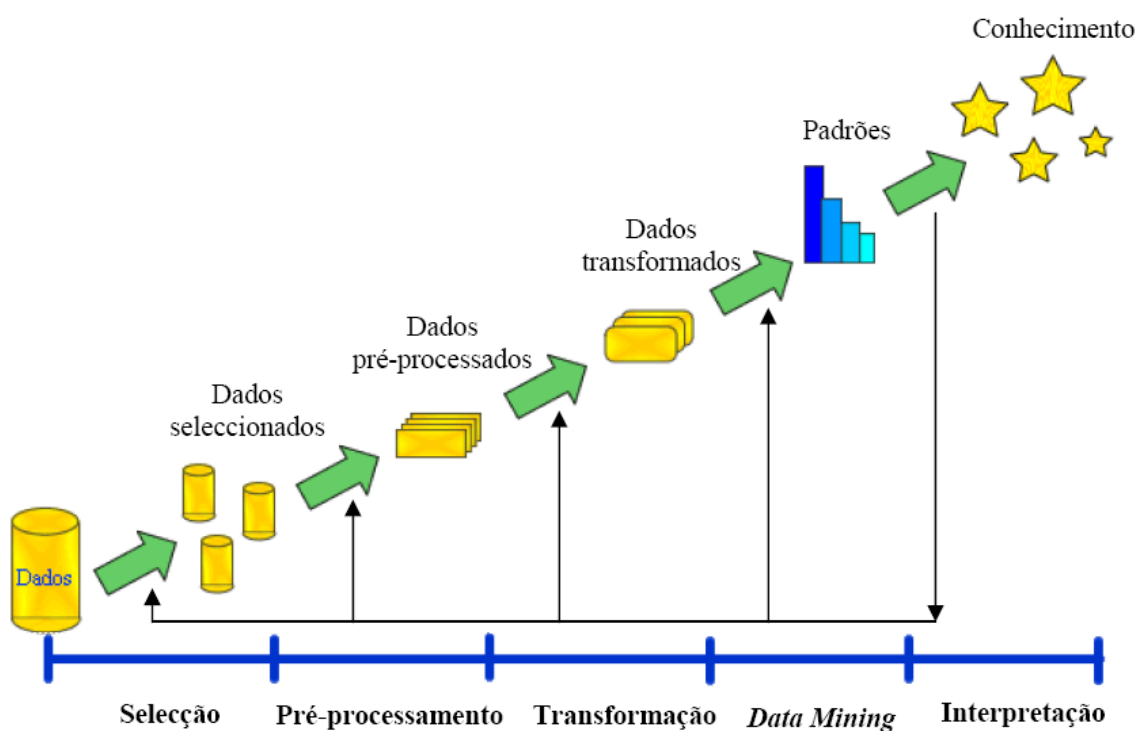


Figura 2 – Processo de Descoberta de Conhecimento em Bases de Dados. Adaptado de Santos e Azevedo [Santos & Azevedo, 2005]

Tabela 2 – Etapas do processo de Descoberta de Conhecimento em Bases de Dados

Etapas do processo de DCBD	Descrição
Seleção	Consiste no estudo/compreensão do domínio de aplicação e na selecção dos dados a analisar. Procura-se, assim, limitar o espaço de pesquisa, direccionando o foco para subconjuntos de variáveis ou dados, onde será realizada a descoberta de conhecimento.
Pré-Processamento	Comporta a limpeza dos dados (<i>data cleaning</i>) de forma a adequar a BD aos algoritmos que, posteriormente, serão utilizados. Inclui, ainda, a integração de dados heterogéneos e a eliminação de dados incompletos/inconsistentes. São estabelecidas estratégias para eliminar ou minimizar os problemas relacionados com dados omissos/incompletos, com a presença de ruído e de erros. Quando necessário, pode ocorrer, a reconfiguração dos dados para assegurar uma BD uniforme e consistente.
Transformação	Visa armazenar os dados num formato que facilite a sua utilização por parte das ferramentas de <i>Data Mining</i> .
<i>Data Mining</i>	Envolve a selecção e utilização dos métodos e técnicas que melhor se enquadrem no cumprimento das metas inicialmente estabelecidas. Depende, fundamentalmente, dos objectivos do processo de <i>Data Mining</i> .
Interpretação	Construção de uma representação do conhecimento acessível e clara. Os padrões identificados pelo sistema são interpretados como conhecimento, o qual pode ser usado para o suporte à tomada de decisão. Este conhecimento traduz-se, por exemplo, no entendimento dos conteúdos da BD ou na explicação de um fenómeno observado.

No início de qualquer projecto de Descoberta de Conhecimento em Bases de Dados é necessário delimitar o domínio sobre o qual se pretende executar o estudo, definir os objectivos a atingir, assim como proceder à recolha de dados e/ou variáveis necessárias à construção da Base de Dados sobre a qual se irá trabalhar. Um outro aspecto que importa destacar prende-se com o estudo do domínio de aplicação, que se constitui

como o fio condutor do desenvolvimento de qualquer projecto de Descoberta de Conhecimento em Bases de Dados [Gopalan & Sivaselvan, 2009].

Antes de se poder dar o processo de Descoberta de Conhecimento em Bases de Dados como concluído deve fazer-se a verificação da qualidade da informação obtida. Esta verificação pode ser vista como a avaliação do desempenho do mecanismo de aprendizagem, sendo crucial no estabelecimento da confiança associada aos resultados obtidos.

A Figura 2, que traduz simbolicamente o processo de Descoberta de Conhecimento em Bases de Dados, pode sugerir a existência de um percurso linear, no entanto, isso pode não se verificar uma vez que, numa dada etapa, pode ser identificada a necessidade de retorno para uma das etapas anteriores. Na verdade, o processo de Descoberta de Conhecimento em Bases de Dados é um processo iterativo e interactivo. Iterativo porque, como foi referido, pode haver necessidade de voltar a fases anteriores. Interactivo porque requer a participação do utilizador sempre que é necessária a tomada de decisões.

Num nível de abstracção mais elevado, as etapas do processo de Descoberta de Conhecimento em Bases de Dados apresentadas na Figura 2 podem ser agrupadas em três tarefas distintas que se apresentam na Figura 3.

A primeira tarefa, denominada pré-processamento dos dados, está directamente relacionada com a adequação dos dados, tendo em vista os modelos e técnicas de *Data Mining* a serem utilizados. A segunda tarefa, designada por *Data Mining*, é o núcleo do processo da Descoberta de Conhecimento em Bases de Dados sendo, muitas vezes, confundida com ele.

Por fim, a terceira tarefa, denominada pós-processamento de dados inclui a visualização, a interpretação e a avaliação dos resultados.

Em projectos de Descoberta de Conhecimento em Bases de Dados é comum atribuir 80% do tempo dispendido à tarefa de pré-processamento dos dados, enquanto os restantes 20% referem-se às restantes tarefas [Adriaans & Zantinge, 1996].

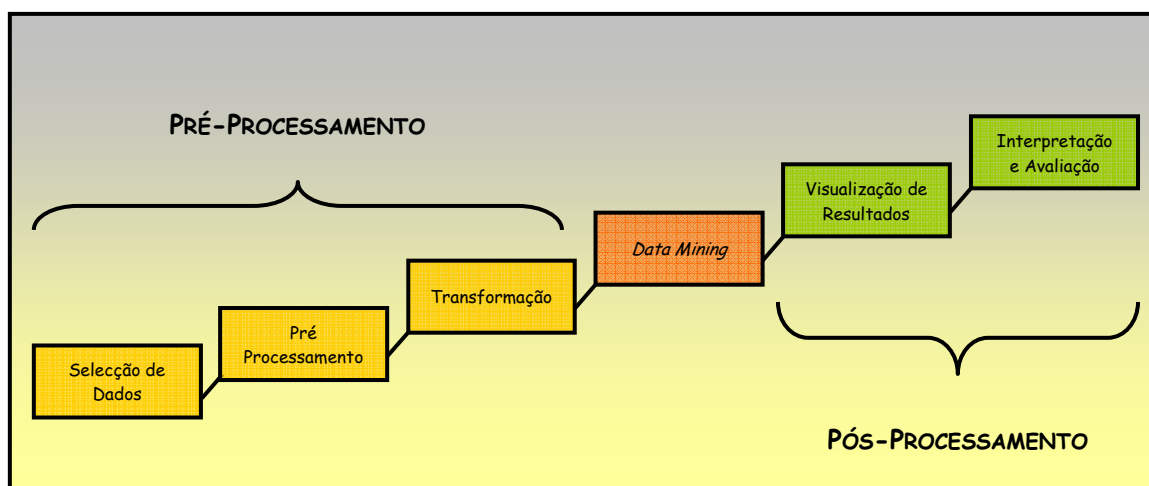


Figura 3 – Etapas do desenvolvimento do processo de DCBD agrupadas em tarefas.

III.2 – *Data Mining*

O termo *Data Mining* pode ser referido através de outros sinónimos como extracção de conhecimento, arqueologia de dados (*data archeology*), colheita de informações (*information harvesting*) e *data dredging* [Fayyad et al., 1996a]. É possível encontrar várias definições para o termo *Data Mining* consoante o autor, a área de especialização e o tipo de abordagem. A Tabela 3 sintetiza algumas dessas definições. A análise dessas definições mostra pontos comuns a todas elas que se podem resumir na seguinte definição: Aplicação de métodos e técnicas em BD que buscam relações, tendências ou regularidades com o intuito de extrair conhecimento implícito nos dados.

A principal diferença entre as ferramentas clássicas de análise de dados e as ferramentas de DM reside na forma como exploram as relações entre os dados. As primeiras utilizam métodos baseados na verificação, i.e. cabe ao utilizador construir hipóteses sobre relações específicas entre as variáveis que vão, através das ferramentas, sendo por ele corroboradas ou refutadas. No que respeita às ferramentas DM são utilizados métodos de descoberta ficando, elas próprias, responsáveis pela procura de relações entre variáveis. Deste modo, não estão dependentes da maior ou menor habilidade do utilizador para propor hipóteses interessantes, para manipular a complexidade dos atributos e para refinar a análise.

Tabela 3 – Algumas das definições mais usuais para o termo *Data Mining*

Definição	Referência
” <i>Data Mining</i> é o processo de encontrar padrões ou relações, previamente desconhecidas e potencialmente úteis, em BD de grande dimensão.”	[Fayyad et al., 1996a]
”O termo <i>Data Mining</i> é vulgarmente utilizado como sinónimo do processo de extracção de informação das BD. Está relacionado com os meios algoritmos através dos quais são extraídos padrões e enumerados os dados.”	[Fayyad, 1997]
” <i>Data Mining</i> refere-se à aplicação de métodos de aprendizagem automática para a enumeração de padrões sobre dados.”	[Michalski et al., 1998]
” <i>Data Mining</i> é o processo de extrair informação ou conhecimento de conjuntos de dados tendo em vista a tomada de decisão.”	[Chang et al., 2001]
” <i>Data Mining</i> é a procura padrões de interesse numa determinada forma de representação, ou conjunto de representações: classificação, árvores de decisão, regras de indução, regressão, segmentação, e assim por diante.”	[Sousa, 2004]
” <i>Data Mining</i> refere-se à utilização de uma variedade de técnicas para identificar “ <i>nuggets</i> ” de informação ou conhecimento, em conjuntos de dados, extraíndo-os de modo a poderem ser utilizados nas áreas de suporte à decisão, previsão e estimativa. Esses dados são, por norma, volumosos e de pouca utilidade quando acedidos de forma directa.”	[Clementine, 2007]
” <i>Data Mining</i> refere-se à utilização de técnicas que procuram estabelecer relações em grandes quantidades de dados, nos mais variados domínios e, na maior parte dos casos, sem supervisão.”	[Cios et al., 2007]

Como foi referido anteriormente, a etapa de DM é apenas uma de várias que compõem o processo de DCBD que, de acordo com Berry e Linoff [Berry & Linoff, 2000] assenta em três pedras basilares, nomeadamente os modelos e as técnicas, os dados e a modelação dos dados. No que concerne aos modelos e técnicas importa destacar que não existe um modelo ou técnica adequada a um determinado problema, podendo ser utilizados vários. A selecção, na maior parte dos casos, é feita com base nos objectivos estabelecidos para o processo de DCBD no geral e, para a etapa de DM no particular. Na Tabela 4 apresentam-se e caracterizam-se alguns desses objectivos, divididos em duas tarefas gerais (previsão e descrição). A previsão tem por objectivo inferir compor-

tamentos futuros com base nas experiências do passado, ao passo que a descrição pretende, de uma forma geral, descrever de forma compacta conjuntos de dados ou associações entre variáveis.

Tabela 4 – Caracterização dos objectivos de *Data Mining*. Adaptado de Cortez [Cortez, 2004]

Objectivo		Descrição
Previsão	Classificação	Corresponde à descoberta de uma função que associa um caso a uma das diversas classes discretas de classificação, de forma a classificar um novo objecto de acordo com um padrão de classificação.
	Regressão	Consiste em prever valores futuros ou desconhecidos de uma variável dependente, a partir de exemplos. Pretende-se conceptualizar um modelo capaz de mimetizar uma função desconhecida, que se aproxime da função dada por um conjunto de características genéricas.
Descrição	Associação ou Dependência	Consiste em encontrar um modelo que descreva dependências significativas entre variáveis, através da identificação de grupo de dados fortemente associados.
	Sumariação	Pretende-se encontrar uma descrição compacta para subconjuntos de dados através de regras de resumo e de descobertas de relações entre as variáveis.
	Segmentação	Permite identificar um conjunto finito de categorias ou segmentos para descrever os dados. Os objectos pertencentes ao mesmo segmento são semelhantes entre si e diferentes dos objectos pertencentes a outros segmentos. Pode ser um objectivo intermédio de DM, sendo realizado numa fase inicial, para encontrar grupos homogéneos de dados para posterior aplicação de algoritmos de aprendizagem.
	Visualização	Trata da apresentação dos resultados de DM através de uma forma visual com o recurso a gráficos ou diagramas. Pretende dar a conhecer informações complexas de uma forma acessível, evidenciando os padrões e as tendências reveladas.

III.3 – Modelos e Técnicas de *Data Mining*

Um modelo pode ser entendido como uma representação matemática, i.e. uma descrição formal e simplificada, de um sistema ou processo com o objectivo de o estudar. Corresponde a uma estruturação da representação do conhecimento, tendo em vista um determinado objectivo que se pretende alcançar. Os modelos podem ser estáticos, quando representam o sistema num determinado instante ou podem ser dinâmicos quando o representam ao longo do tempo, acompanhando as alterações inerentes ao seu funcionamento [Cortez, 2004; Teixeira, 2005]. Um modelo é definido como uma função que atribui a cada exemplo possível, no domínio definido pelos atributos de entrada, um valor contido no domínio dos atributos de saída. Associado a cada modelo está um conjunto de parâmetros que necessitam de ser ajustados ou estimados, a partir de um conjunto de dados, através de um algoritmo de aprendizagem. Após a fase de aprendizagem, é possível utilizar o modelo, i.e. extrapolar novas saídas fornecendo novas entradas [Cortez, 2004; Teixeira, 2005].

Não existe um modelo universal de *Data Mining* que resolva, de forma eficiente, todos os problemas. A escolha de um determinado algoritmo para um dado problema é, de certa forma, uma arte na medida em que é possível utilizar diferentes modelos para a mesma tarefa com vantagens e desvantagens intrínsecas [Fayyad et al., 1996a]. Os modelos e técnicas utilizadas na etapa de *Data Mining* dependem da natureza do problema e dos objectivos que se pretendem atingir. Na Tabela 5 indicam-se algumas das técnicas mais frequentes em função dos objectivos do *Data Mining*.

As técnicas utilizadas no presente trabalho foram as Redes Neurais Artificiais na tarefa de regressão e a Aproximação de Vizinhanças na tarefa de segmentação. Na explicação dos modelos de segmentação, tendo em vista atribuir um novo caso a um determinado segmento, foram utilizadas as Árvores de Decisão. Deste modo, nos subcapítulos subsequentes abordar-se-ão apenas as técnicas de *Data Mining* agora destacadas.

Tabela 5 – Tarefas e técnicas de *Data Mining*. Adaptada de Cortez [Cortez, 2004]

Tarefa Técnica	Previsão		Descrição			
	Classificação	Regressão	Associação	Sumariação	Segmentação	Visualização
Árvores de Decisão	✓	✓		✓	✓	✓
Indução de Regras	✓	✓	✓		✓	
Redes Neurais Artificiais	✓	✓	✓		✓	
Algoritmos Genéticos	✓	✓		✓	✓	
Aproximação de Vizinhanças				✓	✓	

III.3.1 – Árvores de Decisão

A definição mais simples de árvores de decisão é a de uma forma de representação de um conjunto de regras que seguem uma hierarquia de classes ou valores. Expressa uma lógica condicional simples e, do ponto de vista gráfico, são semelhantes a uma árvore.

Uma árvore de decisão representa um conjunto de regras que permite classificar as instâncias quando percorrida desde o nó da raiz até aos nós terminais ou folhas. Cada nó da árvore especifica um teste para os atributos da instância (variáveis) e, cada ramo descendente desse nó, corresponde a um dos valores possíveis para esse atributo. Uma instância é classificada começando por testar o atributo especificado pelo nó raiz e depois seguindo o ramo correspondente ao valor do atributo.

O primeiro nó da árvore de decisão mostra o factor mais correlacionado com os objectos de saída. Os restantes factores são subsequentemente classificados como nós e relacio-

nados com os nós anteriores, possibilitando uma visualização fácil e rápida dos factores que mais se relacionam com os objectos de saída [Quinlan, 1987; Kufrin, 1997; Han & Kamber, 2006; Cios et al., 2007].

As Árvores de Decisão (AD) tiveram a sua origem na área da aprendizagem automática na segunda metade do século XX. Do ponto de vista histórico pode fazer-se referência ao trabalho de Morgan e Sonquist [Morgan & Sonquist, 1963] onde foi proposto um método de criação de ADs, denominado *Automatic Interaction Detection* (AID), com base na separação dos dados em grupos sucessivos. Este método foi posteriormente melhorado com a introdução do teste do chi-quadrado, tendo em vista a identificação de variáveis independentes, dando origem ao algoritmo CHAID. Neste contexto deve ainda salientar-se o algoritmo CART (*Classification and Regression Trees*) [Breiman et al., 1984] e o algoritmo ID3, desenvolvido por Ross Quilan [Quinlan, 1987].

Os algoritmos de indução de árvores de decisão utilizam a aprendizagem supervisionada, i.e. são fornecidas ao sistema as respostas correctas, a partir de um conjunto de exemplos onde, cada um deles, é formado por um vector de entrada e um vector de saída ou resposta [Rich & Knight, 1991].

A filosofia de funcionamento dos algoritmos de indução de AD baseia-se na estratégia “dividir para conquistar”. De um modo geral, baseiam-se na sucessiva divisão do problema em vários sub-problemas, de menor dimensão, até ser encontrada uma solução, i.e. até à obtenção de nós “puros”, em que cada nó represente apenas uma única classe ou a satisfação de um critério de paragem [Rich & Knight, 1991]. As árvores geradas têm a seguinte estrutura [Michalski et al., 1998]:

- Folhas ou nós puros – corresponde às classes/objectos;
- Nós internos – especificam testes efectuados a um único atributo originando duas ou mais sub-árvores que representam as saídas possíveis; e
- Ramos – correspondem aos possíveis valores dos atributos.

Na construção de árvores de decisão levantam-se dois problemas. O primeiro está relacionado com o atributo a seleccionar para teste num determinado nó e, o segundo, está ligado à paragem da divisão dos exemplos. No que respeita à escolha do atributo

mais informativo, existem várias medidas para avaliar a sua capacidade para discriminar as classes. Contudo, todas convergem em dois pontos:

- Uma divisão que mantém as proporções de classes em todas as partições é inútil; e
- Uma divisão onde em cada partição todos os exemplos são da mesma classe tem utilidade máxima [Sousa, 2004; Teixeira, 2005; Han & Kamber, 2006].

As medidas de partição dividem-se em três tipos:

- Medida da diferença dada por uma função baseada nas proporções das classes entre o nó corrente e os nós descendentes, valorizando a pureza das partições, e.g. índice de gini, entropia;
- Medida da diferença dada por uma função baseada nas proporções das classes entre os nós descendentes, que valoriza a disparidade entre as partições; e
- Medida de independência, que mede o grau de associação entre os atributos e a classe.

No que respeita à paragem da divisão dos exemplos, esta deve acontecer quando se verificar uma destas situações [Sousa, 2004; Teixeira, 2005; Han & Kamber, 2006]:

- Todos os exemplos pertencem à mesma classe;
- Todos os exemplos, embora pertencentes a classes diferentes, têm os mesmos valores dos atributos;
- O número de exemplos a dividir é inferior a um certo limite;
- O mérito de todos os possíveis testes de partição dos exemplos é muito baixo.

O algoritmo ID3 [Quinlan, 1987], como foi referido anteriormente, foi um dos algoritmos pioneiros que permitiu o desenvolvimento das árvores de decisão. Desde essa altura têm sido introduzidas melhorias e funcionalidades que resultaram no aparecimento de novas versões/evoluções deste algoritmo como, por exemplo, os algoritmos C4.5 e C5.0.

Uma das limitações exibidas pelo algoritmo ID3 prende-se com a possibilidade de ocorrer sobre-ajustamento, que se manifesta na construção de árvores com elevado desempenho nos dados de treino, mas com um baixo desempenho nos dados de teste.

No algoritmo C4.5 foram introduzidos melhoramentos que permitiram ultrapassar o problema de sobre-ajustamento através da introdução da poda da árvore. A poda da árvore consiste na redução de algumas sub-árvores, transformando-as em folhas, tendo por base a comparação entre os erros nesse nó e a soma dos erros nos nós que dele descendiam. Deste modo consegue-se obter árvores menores, mas com maior capacidade de previsão para novos casos (Quinlan, 1993; Han & Kamber, 2006)

O algoritmo C4.5 permite, ainda, ultrapassar problemas relacionados com valores numéricos, valores omissos, dados com ruído e possibilita a utilização de validação cruzada [Quinlan, 1993].

O algoritmo C5.0 é o sucessor mais recente do algoritmo C4.5, otimizado para lidar com as exigências do mundo real, através do aumento da eficácia ao nível do tempo de processamento e da memória utilizada. Uma das características mais importantes deste algoritmo consiste em gerar vários classificadores, a partir do mesmo conjunto de treino, permitindo depois combiná-los num único classificador final, no qual cada classificador inicial participa com um determinado peso que é ajustado durante a fase de treino. Esta metodologia, designada por “*boosting*”, permite uma redução dos erros de classificação que, nalguns casos, pode atingir cerca de 40% [Quinlan, 1996; Schapire, 2002].

As árvores de decisão são um dos métodos mais simples e adequados para tratar problemas com muitas dimensões. Trata-se de um método preditivo não linear, rápido e que apresenta uma elevada expressividade na representação do conhecimento (Quinlan, 1993; Berry & Linoff, 2000; Han & Kamber, 2006; Gopalan & Sivaselvan, 2009). Na verdade, as árvores de decisão apresentam as seguintes vantagens:

- Método não-paramétrico – não assume nenhuma distribuição particular para os dados, podendo construir modelos para qualquer função desde que o número de exemplos de treino seja suficiente;

- Independência da escala das variáveis – transformações monótonas das variáveis não altera a estrutura da árvore de decisão;
- Facilidade de interpretação – uma decisão complexa é decomposta numa sucessão de decisões elementares permitindo, de forma rápida, determinar quais os factores que mais influenciam a classificação;
- Adaptável a problemas de regressão – para além de se poderem aplicar a problemas de classificação (Árvores de Classificação) podem, também, ser utilizadas em problemas de regressão (Árvores de Regressão);
- Facilidade em lidar com diversos tipos de informação (real, nominal, ordinal);
- Facilidade em lidar com informação incompleta; e
- Escolha automática dos atributos mais relevantes.

As principais desvantagens das árvores de decisão prendem-se com o facto de necessitarem de uma grande quantidade de dados para descobrir estruturas complexas, de apresentarem alguma instabilidade, i.e. pequenas perturbações no conjunto de treino podem provocar grandes alterações no modelo e exibirem fronteiras lineares e perpendiculares aos eixos [Berry & Linoff, 2000; Maimon & Rokach, 2005; Tanianar, 2008; Gopalan & Sivaselvan, 2009].

A indução de regras surge, muitas vezes, associada às árvores de decisão, na medida em que são utilizadas para expressar o conhecimento nelas representado. A regra mais importante corresponde ao primeiro nó da árvore e as seguintes correspondem aos nós subsequentes de acordo com a sua importância [Santos & Azevedo, 2005; Han & Kamber, 2006; Nisbet et al., 2009]. A representação é escrita numa sintaxe do tipo **SE condição ENTÃO acção**, onde a *condição* da regra especifica alguns atributos e valores e a *acção* prevê um valor para um atributo cuja previsão é desejada [Fayyad et al., 1996a]. As regras têm dois graus associados, o suporte e a confiança. O suporte está relacionado com o número de casos no qual a regra é encontrada, enquanto a confiança é a probabilidade condicional da regra (Quinlan, 1993).

III.3.2 – Redes Neurais Artificiais

Um dos objectivos da disciplina de Inteligência Artificial passa pelo desenvolvimento de paradigmas e algoritmos para realizar tarefas cognitivas, as quais são executadas pelos seres humanos de forma eficiente. Neste contexto, inspiradas e influenciadas pelo sistema nervoso dos seres vivos, em particular o do ser humano, surgiram as Redes Neurais Artificiais (RNAs) [Patterson, 1996].

Uma RNA é um conjunto de elementos de processamento simples, designados por neurónios artificiais ou nodos, numa estrutura paralela muito interconectada. Assemelham-se ao comportamento do cérebro uma vez que, por um lado, o conhecimento é adquirido, a partir de um ambiente, através de processos de aprendizagem e, por outro lado, porque o conhecimento é armazenado nas ligações entre os nodos.

Um neurónio (Figura 4) é constituído por um núcleo, por um corpo celular, por um numeroso conjunto de dendrites (destinados a receber sinais de outros neurónios), e por um axónio que pode ou não apresentar ramificações, cuja função é transmitir um estímulo ao(s) neurónio(s) a que está conectado. Os neurónios ligam-se entre si através de junções especializadas, denominadas sinapses, em que uma célula influencia a outra através da passagem de mensagens químicas [Bose & Liang, 1996; Kartalopoulos, 1996; Arbib, 2003].

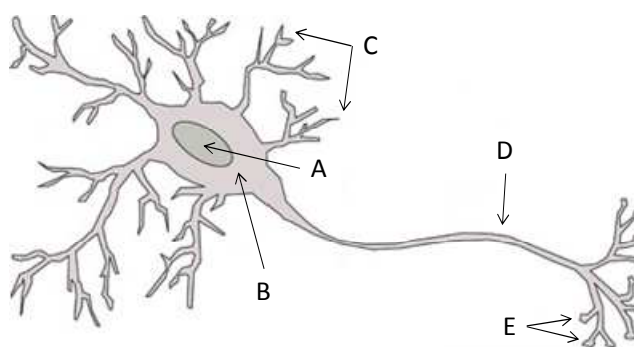


Figura 4 – Representação esquemática de neurónio. A – Núcleo, B – Corpo Celular, C – Dendrites, D – Axónio e E – Botões Terminais

O neurónio que envia a mensagem denomina-se por pré-sináptico e o que a recebe por pós-sináptico. A separá-los existe a fenda sináptica, com cerca de 20 a 40 nm de largura. Os estímulos transmitidos através das sinapses podem ser inibitórios ou excitatórios. Cada neurónio “decide” se irá ou não gerar um potencial de acção após somar todos os estímulos que recebe. Esta “soma” pode ser espacial (quando se está perante estímulos simultâneos de sinapses localizadas em diversos locais da célula pós-sináptica) ou temporal (quando resulta da adição de estímulos de uma mesma sinapse em sequência rápida).

Os neurónios, apesar de processarem a informação a uma velocidade 5 a 6 vezes menor quando comparada com a velocidade de processamento de uma porta lógica de sílicio, compensam essa lentidão pela utilização de uma estrutura maciçamente paralela. Na verdade, os neurónios tendem a agrupar-se em camadas, existindo três tipos principais de conexões:

- Divergentes – onde um neurónio pode estar ligado a outros neurónios através de uma arborização do axónio;
- Convergentes – onde vários neurónios podem estar conectados a um único neurónio; e
- Encadeadas ou cíclicas – as quais podem envolver vários neurónios e formarem ciclos.

Do ponto de vista histórico, as RNAs tiveram a sua origem na década de 40 do século XX, com o trabalho de Warren McCulloch e Walter Pitts. Os referidos autores publicaram um artigo, que se tornou um marco na área científica da Inteligência Artificial, onde apresentaram um modelo simplificado dos neurónios baseado no facto do neurónio, num dado instante, ou estar activo ou inactivo, o que corresponde ao verdadeiro/falso da lógica proporcional ou ao um/zero da álgebra booleana [McCulloch & Pitts, 1943]. Outras contribuições se seguiram, destacando-se a de Hebb [Hebb, 1949] e a de Rosenblatt [Rosenblatt, 1958] que introduziram, respectivamente, o modelo básico de rede de auto-organização e o modelo perceptrão.

Os referidos avanços geraram grande euforia na comunidade científica e atraíram para a área diversos investigadores, passando-se a acreditar que as RNAs podiam resolver qualquer problema [Hagan et al., 1996]. Contudo, alguns resultados menos positivos, a falta de algoritmos de aprendizagem e um estudo onde foi formalmente provado que uma rede formada por uma única camada, independente do algoritmo de aprendizagem, é incapaz de resolver problemas linearmente não separáveis (e.g. a função lógica OU exclusivo) [Minsky & Papert, 1969] tiveram efeitos devastadores, relegando a abordagem conexionista para segundo plano.

No início da década de oitenta, devido a diversas contribuições, houve um ressurgimento do interesse pelas RNAs. Das referidas contribuições destacam-se as redes neuronais recorrentes com conexões simétricas [Hopfield, 1982], as redes não supervisionadas do tipo Kohonen [Kohonen, 1982], o aparecimento do algoritmo da retro-propagação [Rumelhart et al., 1986] (que se tornou o algoritmo de treino mais utilizado em redes neuronais unidireccionais), o desenvolvimento de redes denominadas *Radial Basis-Functions* como alternativa às redes neuronais unidireccionais [Broomhead & Lowe, 1988] e o aparecimento de uma classe de redes supervisionadas denominadas Máquinas de Vectors de Suporte (*Support Vector Machines*) para regressão e reconhecimento de padrões [Boser et al., 1992].

O neurónio artificial, comumente designado por nodo, é a unidade de processamento básica de uma RNA. Embora existam diversos tipos de nodos, de um modo geral, todos eles se comportam como comparadores que produzem uma saída quando o efeito cumulativo das entradas excede um dado valor limite. De acordo com Rojas [Rojas, 1996], um nodo pode ser esquematicamente representado como mostra a Figura 5, sendo constituído por três elementos fundamentais:

- Um conjunto de conexões (x_1, x_2, \dots, x_n), cada uma etiquetada por peso (w_{ij}), i.e. um número real ou binário que tem um efeito excitatório ou inibitório consoante apresente valores positivos ou negativos. Por vezes existe uma entrada adicional, denominada “*bias*”, cujo valor é +1 que visa estabelecer as condições operacionais correctas para o nodo;

- Um integrador, que reduz os n argumentos de entrada (estímulos) a um único valor (μ_i). Frequentemente é utilizada a função adição, pesando todas as entradas numa combinação linear; e
- Uma função de activação (f), que pode condicionar a amplitude do sinal de saída (s_i), introduzindo uma componente de não linearidade no processo computacional.

Em termos formais um nodo é descrito pelas seguintes equações:

$$\mu_i = \sum_{j=1}^n (x_j w_{ij}) + bias \quad (1)$$

$$s_i = f(u_i) \quad (2)$$

para o nodo i com n entradas e uma saída, onde μ_i representa o ganho do nodo e s_i denota a saída do nodo.

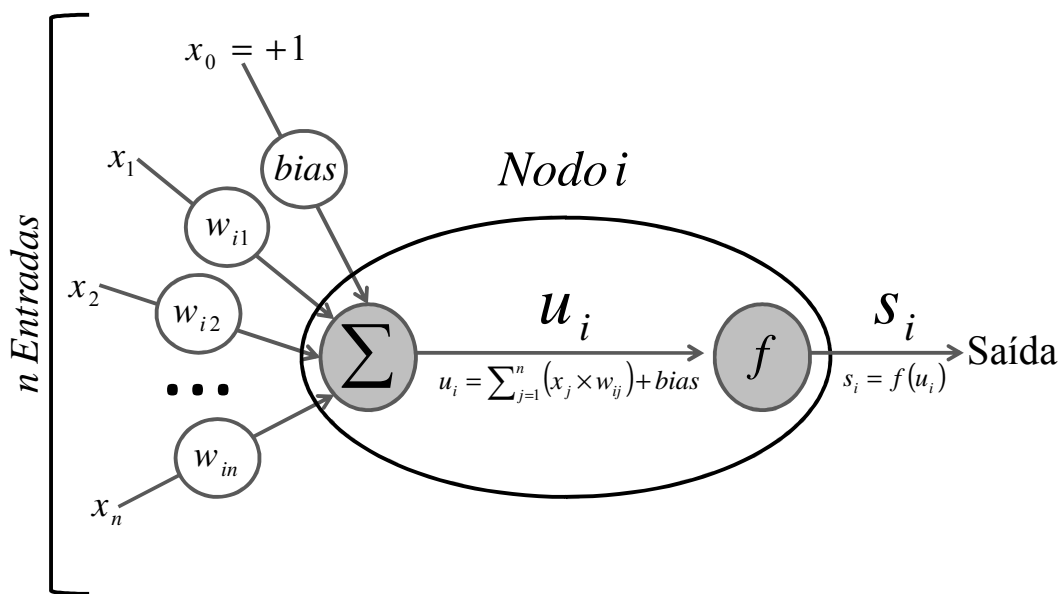
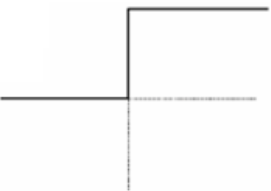
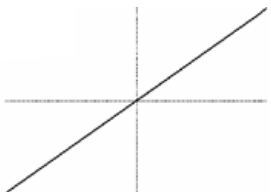

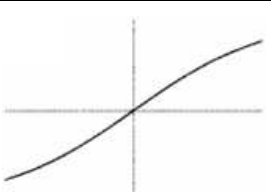


Figura 5 – Estrutura geral de um nodo

A Tabela 6 mostra algumas das funções de activação mais utilizadas. A função limiar, também designada por função *Heaviside*, é normalmente utilizada em nodos do tipo McCulloch-Pitts, em que a saída toma o valor +1 apenas se o ganho for não-negativo, obedecendo a uma filosofia de tudo ou nada. As funções de activação logística e tangente hiperbólica, tal como a gaussiana e a quadrada, introduzem uma componente de não-linearidade ao processo computacional. A função logística, também conhecida por função *sigmoid*, cuja forma pode ser aproximada por um “S”, é a função de activação mais utilizada em RNAs. É uma função crescente que exhibe um balanceamento entre o comportamento linear e não linear, tendendo para a função limiar, no limite, quando k se aproxima de infinito [Jordan, 1995].

Tabela 6 – Algumas funções de activação

Nome	Função	Representação Gráfica
Limiar	$\begin{cases} 1 & , \quad u_i \geq 0 \\ 0 & , \quad u_i < 0 \end{cases}$	
Linear	u_i	
Logística	$\frac{1}{1 + e^{-k u_i}}$	
Tangente Hiperbólica	$\tanh(k u_i)$	

A arquitectura de uma RNA é definida com base no modo como os nodos se interligam na estrutura da rede. Existem inúmeros tipos de arquitecturas de rede, cada um com potencialidades e limitações próprias. De uma forma geral, podem ser divididas em duas categorias principais [Haykin, 2008]:

- Redes Neurais Unidireccionais (RNUs) – são organizadas por camadas, as conexões propagam-se sempre numa só direcção, não existindo ciclos. A mais simples é composta por uma camada de entrada e por uma camada de saída. Contudo, podem possuir uma ou mais camadas internas, cujos nodos são designados por nodos intermédios (Figura 6 (a)); e
- Redes Neurais Recorrentes (RNRs) – A recorrência ocorre, em sistemas dinâmicos, quando a saída de um elemento influencia, de algum modo, a entrada desse mesmo elemento criando, assim, um ou mais circuitos fechados (Figura 6 (b)). Ao conter ciclos, as saídas não dependem apenas das conexões entre nodos mas, também, de uma dimensão temporal. Está-se perante um processo recursivo, obedecendo a condições de paragem, sendo a saída do nodo o valor da última iteração [Rojas, 1996].

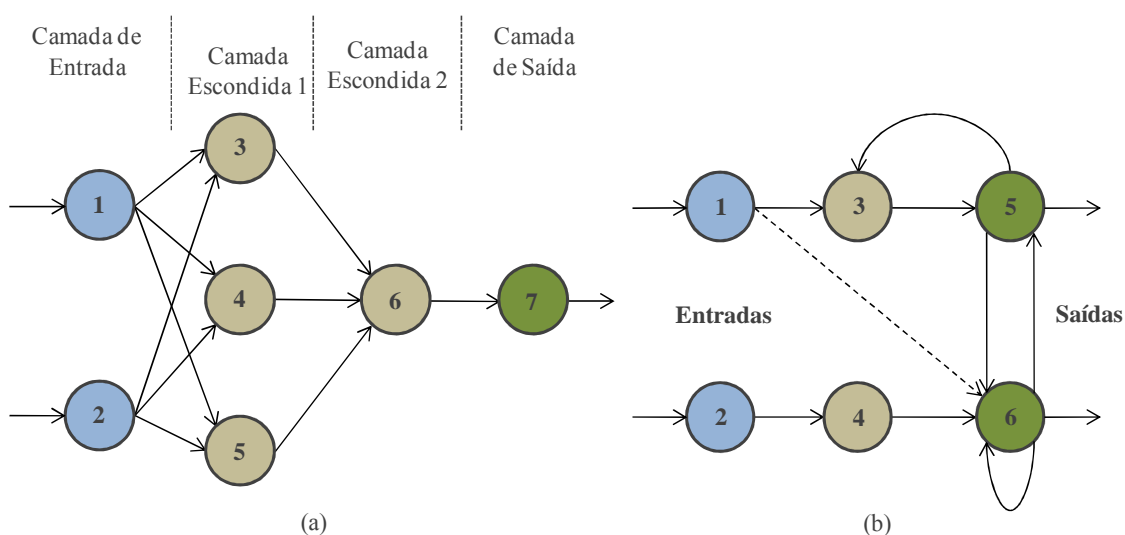


Figura 6 – Arquitectura Unidireccional (a) e Recorrente (b)

Muito embora tenha sido provado que uma camada intermédia é suficiente para apreender qualquer função contínua [Hornik et al., 1989], outros estudos apontam que, na prática, muitas funções são difíceis de apreender com apenas uma camada intermédia [Cheng & Titterington, 1994; Flood & Kartam, 1994]. Na verdade, não existem regras claras para estabelecer quer o número de camadas intermédias, quer o número de nodos que as constituem. Na maior parte dos casos, o estabelecimento da melhor arquitectura de rede é feito por tentativa e erro, começando-se por uma estrutura inicial que vai sendo selectivamente modificada de forma a minimizar uma métrica de erro [Witten & Frank, 2005]. De um modo geral, as ferramentas disponíveis permitem escolher o modo como as redes são construídas, podendo começar-se com redes de pequena/grande dimensão às quais vão sendo, sucessivamente, adicionados/removidos nodos e/ou camadas até ser atingida uma determinada condição de paragem. Muitas ferramentas permitem, ainda, o treino paralelo de várias topologias sendo escolhida aquela que melhor cumpra um determinado critério, e.g. a minimização da média do quadrado dos erros [Witten & Frank, 2005; Clementine, 2007].

Relativamente ao treino de redes neuronais, o algoritmo de retro-propagação [Rumelhart et al., 1986] é tido como o algoritmo de treino mais utilizado em RNUs [Govindaraju & Rao, 2010]. Trata-se de um algoritmo de referência, sendo um método de computação eficiente para o treino de RNUs. Baseia-se em métodos de gradiente descendente e procura minimizar a função de erro no espaço de procura dos pesos, tendo em vista encontrar o conjunto óptimo de valores de pesos que, numa situação ideal, originaria as saídas correctas para uma qualquer entrada.

O algoritmo de retro-propagação envolve essencialmente duas etapas. Na primeira, denominada “propagação frontal”, são fornecidos os vectores de entrada que se propagam entre frente, camada após camada e, com base nas equações (1) e (2), são calculadas as saídas. No final é calculado o erro entre os valores calculados e os valores reais. Na segunda etapa, denominada “retro-propagação”, o erro calculado anteriormente é propagado para trás, desde a camada de saída até à camada de entrada sendo os pesos das conexões iterativamente ajustados [Bose & Liang, 1996].

Antes de se iniciar o treino procede-se à escolha dos valores iniciais dos pesos. Em geral, estes devem ser gerados de forma aleatória, dentro do intervalo $[-0,7; 0,7]$ [Hastie et al., 2001]. Inicia-se então o treino da rede sendo os pesos iterativamente ajustados. O

processo é dado como terminado pela aplicação de critérios de paragem (e.g. quando as alterações nos pesos e no erro forem insignificantes ou até que se atinja um erro mínimo). Uma iteração considera-se terminada quando todos os exemplos de treino tiverem sido considerados.

O algoritmo de retro-propagação é controlado por dois parâmetros, o coeficiente de momento (μ) e taxa de aprendizagem (η), cujos valores se situam no intervalo $[0, 1]$. O coeficiente de momento é utilizado na fase de actualização dos pesos e visa garantir que esta seja feita de um modo consistente. Comporta-se como uma memória, tendo um efeito estabilizador em situações oscilatórias. A taxa de aprendizagem, por seu lado, controla as mudanças nos pesos. Quando η toma valores pequenos, próximos de zero, as mudanças produzidas são pequenas e podem tornar a aprendizagem demasiado lenta. No caso de η ter valores próximos da unidade as variações são demasiado grandes e podem conduzir a instabilidade no treino [Han & Kamber, 2006; Haykin, 2008].

O poder computacional das RNAs assenta em dois aspectos fundamentais. Por um lado, uma topologia que premeia o paralelismo e, por outro, a capacidade de aprendizagem e generalização [Groth, 2000; Nisbet et al., 2009]. Contudo, as RNAs não podem ser tomadas como “caixas mágicas” capazes de dar resposta a todos os problemas. Não raras vezes, necessitam de ser integradas em outros sistemas ou novos paradigmas de resolução de problemas [Haykin, 2008] apresentando, no entanto, características únicas que não se encontram noutros mecanismos ou técnicas [Berry & Linoff, 2000]:

- Aprendizagem e generalização – conseguem descrever o todo a partir de algumas partes, constituindo-se como formas eficientes de aprendizagem e armazenamento de conhecimento;
- Processamento maciçamente paralelo – permite que tarefas complexas sejam realizadas num curto espaço de tempo;
- Não linearidade – consegue dar resposta a problemas reais cuja natureza é não linear;
- Adaptabilidade – pode adaptar a sua topologia de acordo com mudanças de ambiente;

- Robustez e degradação suave – permitem processar o ruído ou informação incompleta/contraditória de forma eficiente, assim como são capazes de manter o seu desempenho quando ocorre desactivação de algumas das suas conexões e/ou nodos;
- Flexibilidade – apresenta um grande domínio de aplicabilidade; e
- Usabilidade – não necessita de conhecimento explícito acerca da função a ser aprendida.

III.3.3 – Aproximação de Vizinhanças

A técnica de aproximação de vizinhanças é baseada no princípio de que registos semelhantes estão próximos uns dos outros, quando analisados numa perspectiva espacial [Perlovsky, 2001; Han & Kamber, 2006; Bramer, 2007; Butenko et al., 2009]. A verificação da localização dos registos, interpretados como pontos no espaço, permite a identificação de regiões, denominadas classes ou segmentos, que definem características comuns para os registos aí incluídos. A segmentação pode ser hierárquica (quando as classes estão organizadas numa hierarquia como, por exemplo, nos dendogramas) ou não hierárquica (quando, pelo contrário, não existe organização entre as classes) [Han & Kamber, 2006; Bramer, 2007].

As técnicas de segmentação não hierárquica, como a estratégia *k-means*, assentam na construção de partições dos objectos armazenados numa BD, num conjunto de k classes, sendo k um parâmetro de entrada. O processo de obtenção das classes é iniciado com k centros em posições aleatórias, as quais são optimizadas iterativamente. Inicialmente cada objecto é atribuído ao centro mais próximo, sendo a distância Euclideana a métrica habitualmente utilizada:

$$d(i, j) = \sqrt{(X_{i1} - X_{j1})^2 + (X_{i2} - X_{j2})^2 + \dots + (X_{in} - X_{jn})^2} \quad (3)$$

onde $i = (X_{i1}, X_{i2}, \dots, X_{in})$ denota o centro do segmento i e $j = (X_{j1}, X_{j2}, \dots, X_{jn})$ denota um objecto contido na BD.

No passo seguinte o algoritmo calcula o ponto médio de cada classe, repetindo-se o processo até não haver alterações ou o número de iterações exceda um determinado limite, previamente definido. De um modo geral, é utilizado, como critério de paragem, a minimização da soma do quadrado dos erros, definida por:

$$E = \sum_{i=1}^k \sum_{p \in C_i} |p - m_i|^2 \quad (4)$$

onde E denota a soma do quadrado dos erros, p denota um objecto da BD e m_i denota o centro da classe C_i [Bradley & Fayyad, 1998].

Cada classe é representada através da localização definida pela média de todos os membros do segmento, comumente designada como centro de gravidade [Witten & Frank, 2005; Han & Kamber, 2006].

Neste tipo de abordagem é utilizada a aprendizagem não supervisionada uma vez que não é fornecido ao sistema qualquer indicação sobre as respostas correctas, sendo a avaliação feita pela utilidade dos agrupamentos formados.

A principal vantagem da estratégia *k-means* reside na simplicidade e na velocidade de processamento, permitindo ser utilizada em BD de grande dimensionalidade. As principais desvantagens prendem-se com o facto de a solução depender da escolha inicial das posições dos centros e utilizar o conceito de média que, em algumas situações, pode não existir. Nestes casos pode utilizar-se a variante *k-medoid*. Nesta técnica o centro de gravidade é atribuído ao objecto da classe que está mais próximo do centro [Han & Kamber, 2006; Chakrabarti et al., 2009; Gopalan & Sivaselvan, 2009].

III.4 – Avaliação dos Modelos

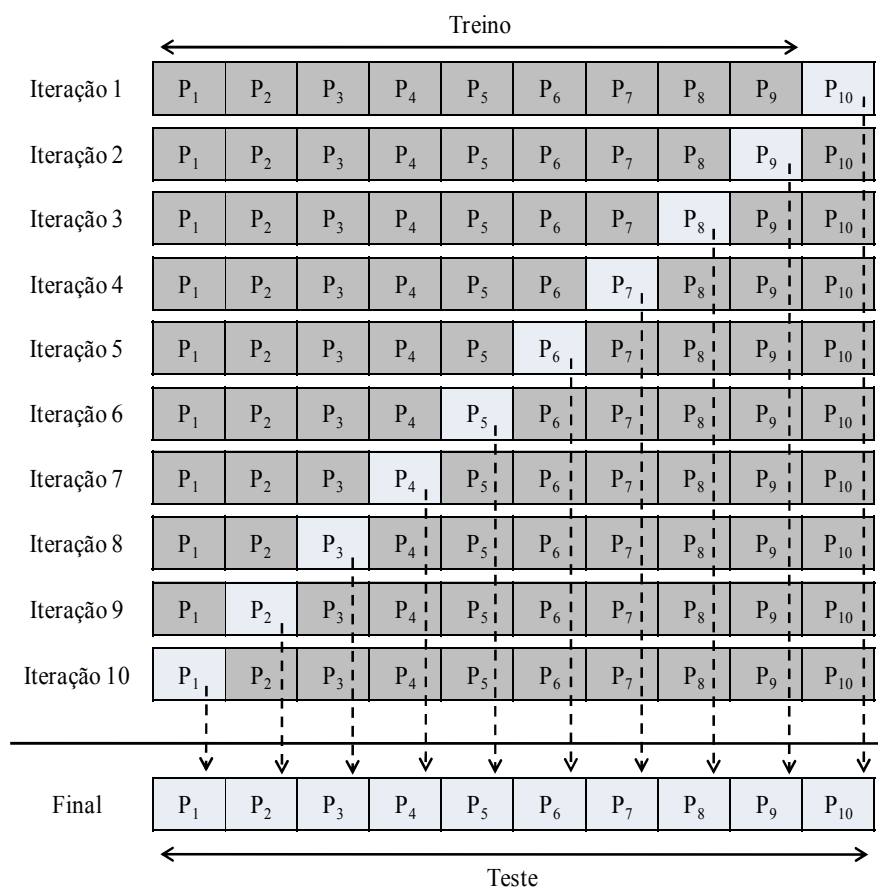
Após a construção dos modelos é necessário avaliar o seu desempenho. Na verdade, esta é uma etapa bastante importante, na medida em que permite inferir qual a confiança que se pode depositar na previsão de cenários futuros. Para conhecer com exactidão a capacidade de generalização de um modelo seria necessário testá-lo em todos os casos possíveis. Contudo, na maioria das aplicações, isso é impraticável uma vez que apenas

se tem disponível um número reduzido de exemplos. Deste modo, ter-se-á que estimar o erro associado ao modelo, tendo o cuidado de não efectuar uma estimativa demasiado optimista ou demasiado pessimista.

Existem diversos métodos para estimar a capacidade de generalização de um modelo, como sejam a estatística simples, a validação por resubstituição, a validação com divisão da amostra, a validação cruzada e o *bootstrapping* [Weiss & Kulikowski, 1990; Han & Kamber, 2006; Gopalan & Sivaselvan, 2009]. No caso da validação por resubstituição, após a construção do modelo, este é utilizado para prever os exemplos utilizados na sua criação sendo estimado o erro associado [Breiman et al., 1984]. A principal desvantagem deste método está relacionada com o facto de a minimização do erro estar subjacente à própria construção dos modelos. Deste modo, a estimativa do erro pode ser demasiado optimista.

A validação com divisão da amostra é um dos métodos mais utilizados para estimar os erros de generalização. Baseia-se na divisão dos dados do problema em dois subconjuntos, mutuamente exclusivos, um para gerar o modelo, designado como conjunto de treino, e outro para a sua validação, denominado conjunto de teste. A cardinalidade dos referidos subconjuntos é variável e depende, entre outros factores, da dimensão da base de dados. Este método apresenta como vantagens a sua simplicidade e rapidez e, como desvantagem, a redução efectiva dos dados disponíveis para o treino [Hand et al., 2001; Larose, 2006; Williams & Simoff, 2006; Bramer, 2007].

Uma variante à validação com divisão de amostra é a validação cruzada k -desdobrável. Neste método os dados (P) são divididos em k subconjuntos mutuamente exclusivos (P_1, P_2, \dots, P_k) de cardinalidade semelhante. Os valores de k podem variar, embora o valor dez seja o mais utilizado. Na primeira iteração são utilizados os dados dos subconjuntos P_1, P_2, \dots, P_9 para criar o modelo reservando-se o subconjunto P_{10} para teste. Na segunda iteração utilizam-se os dados dos subconjuntos P_1, P_2, \dots, P_8 e P_{10} para construir o modelo e o subconjunto P_9 para o testar, repetindo-se o processo (Figura 7). Deste modo, os modelos são treinados e testados k vezes sendo o erro final da generalização dado pela média dos erros de validação, obtidos durante os k testes.

Figura 7 – Validação cruzada com dez iterações ($k = 10$)

A principal vantagem está relacionada com a utilização de todos os casos disponíveis tanto para construir o modelo como para o validar. A principal desvantagem, por seu lado, prende-se com o facto de exigir um grande esforço computacional [Han & Kamber, 2006; Larose, 2006; Williams & Simoff, 2006; Bramer, 2007].

A validação cruzada é notavelmente superior à validação com divisão da amostra para Bases de Dados com um número reduzido de registos, particularmente em problemas de classificação, quando é mantida a distribuição percentual das várias classes nas diversas partições (P_1, P_2, \dots, P_k) [Weiss & Kulikowski, 1990].

III.4.1 – Avaliação de modelos de classificação

Quando se trata de problemas de classificação uma das técnicas de avaliação mais utilizadas é a Matriz de Coincidências [Kohavi & Provost, 1998]. Trata-se de uma

matriz $N \times N$ (onde N denota o número de classes possíveis), na qual são confrontados os valores previstos pelo modelo e os valores desejados, sob a forma de tabela de duas entradas. As entradas horizontais referem-se às saídas desejadas, enquanto as saídas devolvidas pelo modelo se encontram nas entradas verticais. As células são preenchidas com os números de casos que correspondem ao cruzamento das entradas. O erro do classificador pode ser definido como sendo a razão entre o número de casos incorrectamente classificados e o número total de casos.

Na Tabela 7 apresenta-se um exemplo de uma matriz de coincidências onde é possível visualizar, de um modo rápido e inequívoco, o desempenho do modelo. Neste caso, relativamente à classe X , o modelo classificou correctamente 45 instâncias e incorrectamente 4 (três classificadas como Y e uma como Z). No que concerne à classe Y todas as instâncias foram correctamente classificadas e, em relação à classe Z , foram classificadas correctamente 32 instâncias tendo apenas uma sido classificada incorrectamente. O erro do classificador é, neste caso, 4,9 % (5 instâncias mal classificadas em 128).

Um outro problema que se pode colocar na avaliação do desempenho de um modelo de classificação prende-se com os custos do erro. Na verdade, em algumas aplicações existem diferentes custos associados aos vários erros possíveis. Por exemplo, se está em causa a classificação de uma massa de água superficial de acordo com as suas características de qualidade para usos múltiplos [SNIRH, 2010], é muito mais gravoso, do ponto de vista da saúde pública, classificar uma massa de água poluída como não poluída do que o contrário.

Tabela 7 – Matriz de coincidências 3×3

Classe	X	Y	Z
X	45	3	1
Y	0	46	0
Z	0	1	32

O custo do erro pode ser tomado como uma penalização imposta ao modelo no caso de este cometer um determinado erro. Quando se pretende fazer a avaliação do modelo com base nos custos da classificação ter-se-á que conhecer as penalizações a atribuir, definidas, em geral, através de uma matriz de custos. Trata-se de uma matriz em tudo semelhante à matriz de coincidências, sendo as células preenchidas com o custo de cada tipo de erro, Tabela 8.

O custo do classificador é calculado de acordo com a equação:

$$Custo = \sum_{i=1}^N \sum_{j=1}^N C_{i,j} M_{i,j} \quad (5)$$

onde $C_{i,j}$ e $M_{i,j}$ denotam, respectivamente, os valores da linha i , coluna j da matriz de custos e da matriz de coincidências. Para o exemplo apresentado na Tabela 7 e na Tabela 8 o custo será:

$$Custo = 0 \times 45 + 3 \times 3 + 10 \times 1 + 2 \times 0 + 0 \times 46 + 12 \times 0 + \\ + 8 \times 0 + 1 \times 1 + 0 \times 32 = 20 \quad (6)$$

O custo médio por decisão, definido em termos da razão entre o custo da classificação e o número total de casos, é 0,16 (20/128). No entanto, é de salientar que, até aqui, apenas foi tomado em consideração os custos dos erros não se tendo em conta o ganho relativo às classificações correctas. Este factor pode ser incluído na matriz de custos codificando o ganho das classificações correctas com valores negativos.

Tabela 8 – Matriz de custos para um problema de classificação com três classes

Classe	X	Y	Z
X	0	3	10
Y	2	0	12
Z	8	1	0

III.4.2 – Avaliação de modelos de regressão

Um modelo de regressão tenta estimar o valor associado a cada um dos exemplos. O principal objectivo de um método de regressão é conceber o “melhor” modelo na perspectiva da minimização de uma medida do erro global. Define-se por erro ou resíduo a diferença entre o valor observado e o valor previsto, podendo ser calculado um erro global, respeitante a todos os valores previstos, usando diversas medidas, nomeadamente o Erro Absoluto Médio (EAM), a Soma do Quadrado dos Erros (SQE), a Média do Quadrado dos Erros (MQE) ou a Raiz Quadrada da Média do Quadrado dos Erros (RQMQE), definidas através das equações [Witten & Frank, 2005]:

$$EAM = \frac{\sum_{i=1}^N |e_i|}{N} \quad (7)$$

$$SQE = \sum_{i=1}^N e_i^2 \quad (8)$$

$$MQE = \frac{SQE}{N} \quad (9)$$

$$RQMQE = \sqrt{MQE} \quad (10)$$

onde $e_i = Y_i - \hat{Y}_i$ denota o erro ou resíduo, Y_i e \hat{Y}_i denotam, respectivamente o valor observado e o valor previsto pelo modelo e N denota o número de observações.

O desempenho de um modelo de regressão será tanto melhor quanto menor for o erro global. As medidas EAM e MQE são duas das métricas mais utilizadas no âmbito da avaliação dos modelos de regressão. Ambas têm as suas origens na estatística, tendo como intuito definir quão melhor é um estimador para um conjunto de valores. Segundo Torgo [Torgo, 1999] estas medidas, quando aplicadas à avaliação de modelos de regressão, servem diferentes propósitos. No caso de ser aceitável cometer alguns erros extremos, desde que o modelo se aproxime, na maior parte das vezes, do valor

observado, a minimização do EAM é a métrica mais adequada. Quando, pelo contrário, se está perante situações em que é crucial não cometer erros extremos, a MQE deve ser minimizada uma vez que esta métrica, ao utilizar o quadrado do erro, amplifica os erros extremos.

Muito embora as métricas apresentadas anteriormente possibilitem avaliar o desempenho dos modelos de regressão, elas não fornecem informação sobre se, de um modo geral, o modelo apresenta tendência para subestimar ou sobrestimar as variáveis dependentes. O cálculo da média dos erros, denominada *bias* (equação 11), pode fornecer a resposta à referida questão [Chenard & Caissie, 2008]. Atendendo a que $e_i = Y_i - \hat{Y}_i$ onde Y_i denota o valor observado e \hat{Y}_i denota o valor previsto pelo modelo, valores de *bias* negativos significam que, em média, o modelo tem tendência para sobrestimar as variáveis dependentes. Valores de *bias* positivos, pelo contrário, significam tendência para subestimar as variáveis de saída.

$$bias = \frac{\sum_{i=1}^N e_i}{N} \quad (11)$$

Uma outra métrica comumente utilizada para avaliar o desempenho dos modelos de regressão é o coeficiente de determinação (R^2). De acordo com Chenard e Caissie [Chenard & Caissie, 2008] e com Singh e seus colaboradores [Singh et al., 2009] o coeficiente de determinação representa a percentagem de variabilidade que pode ser explicada pelo modelo, sendo calculado através da equação:

$$R^2 = \left[\frac{N \sum_{i=1}^N Y_i \hat{Y}_i - \sum_{i=1}^N Y_i \times \sum_{i=1}^N \hat{Y}_i}{\sqrt{N(\sum_{i=1}^N Y_i^2) - (\sum_{i=1}^N Y_i)^2} \times \sqrt{N(\sum_{i=1}^N \hat{Y}_i^2) - (\sum_{i=1}^N \hat{Y}_i)^2}} \right]^2 \quad (12)$$

onde Y_i , \hat{Y}_i e N , tal como anteriormente, denotam o valor observado, o valor previsto pelo modelo e o número de observações.

Neste contexto é, ainda, possível usar métricas relativas como, por exemplo o Erro Absoluto Médio Relativo (EAMR). Estas métricas são obtidas calculando o quociente entre as equações (7), (8), (9) e (10) e equações similares onde, em vez do erro ou

resíduo, e_i , é utilizado o erro em relação à média dos valores de saída, $\bar{e}_i = Y_i - \bar{Y}_i$, onde Y_i e \bar{Y}_i denotam, respectivamente, o valor observado e o valor médio da saída i [Witten & Frank, 2005]:

$$EAMR = \frac{\sum_{i=1}^N |e_i|}{\sum_{i=1}^N |\bar{e}_i|} \quad (13)$$

As principais vantagens das métricas relativas prendem-se, por um lado, com o facto de serem independentes da escala dos valores de saída e, por outro, compararem o desempenho do modelo com uma previsão simples, feita através da média dos valores. A medida relativa do erro pode apresentar valores inferiores, superiores ou iguais à unidade significando que o modelo de previsão apresenta um desempenho, respectivamente, melhor, pior ou similar à previsão via valor médio.

CAPÍTULO IV – PROCEDIMENTO EXPERIMENTAL

A componente experimental deste trabalho divide-se em duas partes. A primeira apresenta um carácter analítico, sendo constituída pela determinação, no laboratório, de alguns parâmetros de qualidade da água. A segunda parte refere-se à construção de uma base de dados, tendo sido utilizado um conjunto de dados que o laboratório já possuía. Escolheu-se trabalhar os parâmetros químicos de qualidade da água dos últimos três anos, de Janeiro de 2007 a Dezembro de 2009. Para o efeito foi necessário extrair os valores do programa que o Laboratório utiliza para emitir os Relatórios de Análise, designado por *LabWay-LIMS*.

Construída a base de dados de qualidade da água para consumo humano relativa a treze dos catorze concelhos do Distrito de Évora, esta vai ser utilizada para o desenvolvimento de modelos baseados no processo DCBD.

IV.1 – Amostragem e Preservação das Amostras

A exactidão dos resultados analíticos depende de vários factores, dos quais se destacam o modo como é efectuada a colheita das amostras, o seu acondicionamento, o transporte até ao laboratório e o seu armazenamento. A amostra deve ser, tanto quanto possível, representativa da água que se deseja analisar i.e., de acordo com Zhang [Zhang, 2007], ter as propriedades médias dessa água, mantendo as proporções relativas de todas as espécies que se pretendem quantificar. Deste modo, torna-se necessário que a colheita das amostras seja feita por técnicos qualificados procedendo conforme as directivas vigentes. No caso das amostras analisadas no Laboratório de Saúde Pública de Évora, a colheita é efectuada por Técnicos de Saúde Ambiental.

A representatividade da amostra pode ser afectada por processos físicos, químicos, e/ou biológicos. Entre os processos físico-químicos que alteram as propriedades das amostras de água durante a recolha destacam-se a degasificação, a volatilização, a precipitação,

a degradação fotofísica, a deterioração da matéria em suspensão, a degradação química, a redistribuição dos analitos, a contaminação e a adsorção.

Um outro aspecto importante, que deve ser tido em consideração, está relacionado com os recipientes utilizados na colheita de amostras. Estes devem preservar a composição da amostra, evitando perdas por adsorção, evaporação ou contaminação por substâncias estranhas. Na escolha do recipiente deve atender-se, ainda, à sua resistência térmica e mecânica, à facilidade em abrir e fechar, ao tamanho, peso e forma. Na Tabela 9 pode-se verificar quais os recipientes utilizados na colheita, o modo de preservação da amostra e o intervalo de tempo máximo entre a realização da amostragem e a determinação analítica. De um modo geral, são seguidas as recomendações do *Standard Methods for the Examination of Water and Wastewater* (SMEWW) [Eaton et al., 2005].

Tabela 9 – Colheita e preservação de amostras de águas

Parâmetro	Recipiente	Execução	Preservação	Armazenamento
Condutividade	Polietileno	*	Refrigerar	28 dias
Ferro	Vidro	-	Acidificar até $1 < \text{pH} < 2$ com H_2SO_4 4,5 M	1 mês
Amónio	Polietileno	*	Acidificar até $1 < \text{pH} < 2$ com H_2SO_4 e refrigerar	21 dias
Metais	Polietileno tratado	-	Acidificar até $1 < \text{pH} < 2$ com HNO_3	6 meses
Nitrato	Polietileno	48 h	Acidificar até $1 < \text{pH} < 2$ com HCl e refrigerar	7 dias
Nitrito	Polietileno	24 h	Refrigerar	2 dias
Oxidabilidade	Polietileno	*	Refrigerar no escuro	2 dias
pH	Polietileno	*	Refrigerar	6 horas
Potássio e Sódio	Polietileno tratado	-	Acidificar até $1 < \text{pH} < 2$ com HNO_3	1 mês
Turvação	Polietileno	*	Refrigerar no escuro	48 horas
Microbiológicos	Polietileno	*	Refrigerar	6 horas

* Tão cedo quanto possível

IV.1.1 – Técnicas de Amostragem

O procedimento técnico para a colheita de amostras varia consoante o tipo de análise a efectuar (microbiológica ou química) e, também, o acesso à amostra. A Recomendação nº 8/2005 do Instituto Regulador de Águas e Resíduos (IRAR) determina a sequência de amostragem. De acordo com o referido documento, em primeiro lugar faz-se a colheita das amostras que obrigatoriamente tenham de ser colhidas sem escoamento prévio (chumbo, cobre e níquel), seguida da colheita destinada às análises microbiológicas e, por fim, a colheita destinada aos restantes parâmetros químicos.

IV.1.1.1 – Colheita de Amostras de Água para Análise de Níquel, Chumbo e Cobre

Abrir a torneira e colher directamente para o recipiente adequado o primeiro litro de água. Salienta-se que, de acordo com a Recomendação nº 8/2005 do IRAR, é obrigatória a recolha de 1 dm³. Opcionalmente, esta amostra pode, também, ser utilizada para análise dos restantes metais.

IV.1.1.2 – Colheita de Amostras de Água para Análises Microbiológicas

Na recolha de amostras de águas para consumo humano, destinadas a análises microbiológicas existe um conjunto de procedimentos que devem ser atendidos. No caso da colheita ser realizada numa torneira deve-se:

- Lavar e desinfectar bem as mãos;
- Retirar qualquer filtro ou mangueira colocado na torneira;
- Abrir a torneira e deixar correr água o tempo suficiente para esgotar a água que tenha estado parada na canalização;
- Fechar a torneira e flamejar bem (no interior e exterior) com maçarico ou tocha embebida em álcool;

- Abrir de novo a torneira e deixar correr a água livremente até arrefecer;
- Abrir o frasco de colheita, previamente esterilizado, e colher a água mantendo-o inclinado, para evitar a sua contaminação pelo ar. Manter a rolha na mão esquerda, virada para baixo, tendo o cuidado de nunca tocar no interior da rolha ou gargalo do frasco;
- Fechar imediatamente o frasco de colheita; e
- Colocar uma etiqueta que contenha a identificação da amostra, assim como o nome do requisitante, ponto de amostragem e data da colheita.

No caso do ponto de amostragem ser um fontanário, procede-se de igual forma, mas sem flamejar a bica.

Um aspecto importante, que deve ser observado, está relacionado com o facto de ser necessário deixar-se uma camada de ar para permitir a homogeneização da amostra no laboratório. Assim, os frascos de colheita não devem ser completamente cheios.

IV.1.1.3 – Colheita de Amostras de Água para Análises Químicas

O recipiente é lavado, previamente, com a própria amostra. Em seguida faz-se a colheita, tendo o cuidado de encher completamente o frasco, Só depois de garantir a ausência do ar no interior do frasco se procede ao seu fecho e à sua identificação.

IV.1.2 – Acondicionamento e Transporte das Amostras

O transporte das amostras até ao Laboratório deve ser feito em malas térmicas, de forma a garantir que a amostra mantenha uma temperatura de $5^{\circ}\text{C} \pm 3^{\circ}\text{C}$. O período de tempo entre o momento da colheita e o início das análises não deverá ser superior a seis horas.

IV.1.3 – Pontos e Frequência de Amostragem

O anexo IX do Decreto-Lei n.º 236/98, de 1 de Agosto, estabelecia a frequência mínima de amostragem e análise de águas para consumo humano, para efeitos de vigilância

sanitária, em função da população servida. Em 2008, com a entrada em vigor do Decreto-Lei n.º 306/2007, de 27 de Agosto, deixa de ser feita menção à frequência de amostragem. Apenas é referido que as acções de vigilância sanitária são realizadas pela Autoridade de Saúde, abrangendo a realização de análises complementares às previstas no Plano de Controlo da Qualidade da Água (PCQA) e a avaliação do risco para a saúde humana da qualidade da água para consumo humano. Deste modo, os pontos e a frequência de amostragem são estabelecidos pela Autoridade de Saúde Concelhia que informa o Laboratório, no final de cada ano, de forma a calendarizar o ano seguinte. Tendo em vista a complementaridade e a monitorização contínua, as colheitas referentes aos procedimentos de vigilância e de controlo são, sempre que possível, efectuadas de forma intercalada.

A Autoridade de Saúde definiu três grupos de parâmetros que se apresentam na Tabela 10. A frequência de amostragem segue, de um modo geral, o previsto no anexo IX do Decreto-Lei n.º 236/98, de 1 de Agosto, sendo maior para o grupo P₁ e menor para o grupo P₃, tendo por base a população servida e a análise do registo histórico. Realça-se, no entanto, o facto de as análises referentes aos parâmetros dos grupos P₂ e P₃ serem mais frequentes no Outono e no Inverno por serem, de alguma forma, influenciados pelas chuvas.

Tabela 10 – Parâmetros incluídos no programa de vigilância sanitária divididos por grupos

	Grupo P ₁	Grupo P ₂	Grupo P ₃
Parâmetros Microbiológicos	Bactérias coliformes <i>Escherichia coli</i> <i>Enterococcus</i>	Grupo P ₁ + <i>Clostridium perfringens</i> Número de colónias a 22 °C Número de colónias a 37 °C	Grupo P ₂
Parâmetros Químicos	pH Condutividade	Grupo P ₁ + Amónio, Ferro, Manganês, Nitrato, Nitrito, Oxidabilidade e Turvação	Grupo P ₂ + Cádmio, Chumbo, Cobre, Crómio, Níquel, Potássio e Sódio

IV.2 – Métodos e Técnicas Analíticas

As técnicas e métodos de ensaio utilizados na quantificação dos diferentes parâmetros químicos apresentam-se na Tabela 11 onde se refere, também, o limite de quantificação do método de ensaio. Na Tabela 12, por seu lado, apresentam-se as técnicas e métodos de ensaio utilizados na quantificação dos diferentes parâmetros microbiológicos.

Tabela 11 – Parâmetros químicos analisados e respectivos métodos de ensaio

Parâmetro	Técnica de Ensaio	Método de Ensaio	Limite de Quantificação
Alumínio	Espectrometria de Absorção Atômica	SMEWW 3113-B	11 µg/dm ³
Cádmio			0,5 µg/dm ³
Chumbo			6 µg/dm ³
Cobre			6 µg/dm ³
Crómio			5 µg/dm ³
Manganês			5 µg/dm ³
Níquel			5 µg/dm ³
Amónio	Espectrometria de Absorção Molecular	ISO 7150-1:1984	0,15 mg/dm ³
Ferro		NP 2202:1996	50 µg/dm ³
Nitrato		SMEWW 4500-NO ₃ ⁻	1,5 mg/dm ³
Nitrito		NP EN 26777:1996	0,012 mg/dm ³
Condutividade	Condutimetria	NP EN 27888:1996	26,0 µS/cm
Oxidabilidade	Titulimetria	NP 731:1969	0,9 mg/dm ³
pH	Potenciometria	SMEWW 4500-H ⁺	---
Potássio	Fotometria de Chama de Emissão	SMEWW 3500-K	0,1 mg/dm ³
Sódio		SMEWW 3500- Na	1 mg/dm ³
Turvação	Turbidimetria	ISO 7027:1999	0,25 UNT

Tabela 12 – Parâmetros microbiológicos analisados e respectivos métodos de ensaio

Parâmetro	Técnica de Ensaio	Método de Ensaio
Bactérias coliformes e <i>Escherichia coli</i>	Filtração	ISO 9308-1:2000
<i>Clostridium perfringens</i>		HPA NHS W5:2005
Enterococos intestinais		ISO 7899-2:2000
Número de colónias a 22 °C	Incorporação	ISO 6222:1999
Número de colónias a 37 °C		

CAPÍTULO V – APRESENTAÇÃO, INTERPRETAÇÃO E ANÁLISE DE RESULTADOS E MODELOS

A crescente importância que se atribui à DCBD, com recurso a técnicas de *Data Mining* tem, como foi referido anteriormente, dado origem ao desenvolvimento de inúmeras aplicações de *software* que procuram dar resposta aos problemas e desafios que se têm colocado sendo, de acordo com um levantamento efectuado por Cruz [Cruz, 2007], conhecidas mais de cento e cinquenta ferramentas diferentes.

No presente trabalho procurou-se uma ferramenta que permitisse implementar RNAs, na tarefa de regressão, possibilitasse a segmentação dos dados e a construção dos respectivos modelos explicativos com recurso a Árvores de Decisão, tendo a escolha recaído na ferramenta WEKA (*Waikato Environment for Knowledge Analysis*).

V.1 – Breve descrição da ferramenta WEKA

A ferramenta *WEKA* permite desenvolver o processo de DCBD, uma vez que disponibiliza diversas formas de preparação de dados, contempla uma série de algoritmos de *Data Mining* e possibilita a utilização de diversos métodos, tendo em vista a validação dos resultados. A ferramenta foi desenvolvida na Universidade de Waikato (Nova Zelândia), faz uso das facilidades disponibilizadas pela programação orientada ao objecto, encontrando-se escrito em linguagem *Java* [Hall et al., 2009].

O *software WEKA* possui interfaces gráficas amigáveis permitindo ao utilizador escolher, logo no menu inicial, uma de quatro possíveis interfaces possibilitando, deste modo, a sua utilização por utilizadores menos familiarizados com a ferramenta. No que concerne aos dados, estes têm que estar num formato próprio do *WEKA* (extensão *.arff*), o que pode ser conseguido facilmente utilizando um conversor, disponibilizado pelo *software*, aplicável para ficheiros do tipo “separado por vírgulas” (extensão *.csv*) ou do tipo “separado por tabs” (extensão *.tsv*).

A ferramenta *WEKA* disponibiliza diversos algoritmos de *Data Mining*, entre os quais se salientam as RNAs e as Árvores de Decisão e a Aproximação de Vizinhanças com recurso à estratégia *k-means*. Relativamente à validação dos resultados permite a utilização de validação com divisão da amostra e de validação cruzada. Além do exposto, permite que o utilizador defina o número de repetições tendo em vista assegurar a significância estatística dos resultados.

V.2 – Apresentação e Preparação dos Dados

A preparação dos dados inclui, como se pode observar na Figura 3, as etapas de selecção, pré-processamento e transformação dos dados. A qualidade dos dados e a forma como se encontram estruturados pode condicionar o sucesso e a utilidade de uma análise inteligente dos mesmos. Na etapa de pré-processamento são realizadas operações que visam melhorar a qualidade dos dados, contribuindo para que a etapa de *Data Mining* seja realizada de forma mais eficiente. Quando se analisa uma BD podem ser identificados diversos tipos de problemas que podem estar relacionados com o tamanho e dimensionamento da BD, a volatilidade dos dados, a ausência de algum atributo relevante, a presença de ruído e com a existência de dados incompletos, omissos ou desconhecidos.

Existem diversas técnicas, com origem na estatística e na matemática, para resolver os problemas relacionados com dados incompletos, omissos ou desconhecidos, tais como os métodos Bayesianos [Gelman et al., 2003] e os métodos de imputação [Little, 1992; Rubin, 1996; Schafer, 1999; Mamedov & Udalov, 2002].

Um dos métodos de imputação mais comuns, em BD com elevado número de registos, consiste na eliminação dos exemplos com dados omissos. No entanto, em algumas situações, este método pode conduzir a alterações significativas na representatividade dos dados relativamente ao universo em estudo e conduzir a amostras muito pequenas.

Os parâmetros microbiológicos foram, à partida, excluídos da BD inicial (Anexo I), uma vez que, segundo o Decreto-Lei nº 306/2007, de 27 de Agosto, o valor paramétrico é zero. No que concerne aos parâmetros químicos, verifica-se que existem registos que apresentam valores omissos. Contudo, a existência de dados em falta está relacionada com a natureza do problema em estudo, uma vez que, devido ao facto de se tratar de

procedimentos de vigilância, apenas se efectuarem as análises quando são solicitadas pela Autoridade de Saúde. Deste modo, optou-se por descartar os registos que apresentavam valores omissos relativos aos parâmetros químicos incluídos no grupo P₁ (pH e condutividade), tendo-se passado dos 817 registos iniciais para 789. Em seguida fez-se a caracterização dos variáveis existentes na BD, em termos estatísticos (Tabela 13).

Tabela 13 – Caracterização estatística das variáveis presentes na base de dados

Parâmetro	Mínimo	Máximo	Média	Desvio Padrão	Registos
Condutividade ($\mu\text{S}/\text{cm}$)	102,0	1220	535,8	189,6	789
pH (Escala Sørensen)	5,8	8,7	7,5	0,3	789
Nitrato (mg/dm^3)	1,6	130	19,1	18,4	196
Oxidabilidade (mg/dm^3)	0,92	7,3	2,1	0,9	73
Manganês (mg/dm^3)	2,0	86	13,8	12,6	69
Turvação (UNT)	0,25	16,6	1,2	2,5	60
Potássio (mg/dm^3)	0,41	9,0	2,6	1,5	52
Sódio (mg/dm^3)	9,0	240	37,2	35,0	52
Cobre ($\mu\text{g}/\text{dm}^3$)	0,006	0,20	0,021	0,035	31
Alumínio ($\mu\text{g}/\text{dm}^3$)	32,0	170	88,1	60,0	9
Ferro ($\mu\text{g}/\text{dm}^3$)	50,0	200,0	85,6	45,6	9
Nitrito (mg/dm^3)	0,013	0,040	0,021	0,013	4
Chumbo ($\mu\text{g}/\text{dm}^3$)	6,0	8,2	7,1	1,6	2
Cádmio ($\mu\text{g}/\text{dm}^3$)	1,6	---	---	---	1
Níquel ($\mu\text{g}/\text{dm}^3$)	9,0	---	---	---	1
Amónio (mg/dm^3)	---	---	---	---	0
Crómio ($\mu\text{g}/\text{dm}^3$)	---	---	---	---	0

A análise da Tabela 13 mostra que o número de registos válidos relativos aos parâmetros amónio, nitrito, ferro, cádmio chumbo, cobre, crómio, níquel e alumínio é muito baixo verificando-se que, na maior parte dos casos, ou não eram pedidas análises ou os valores estavam abaixo do limite de quantificação dos métodos utilizados na sua determinação (Tabela 14). Deste modo, optou-se por retirá-los do estudo, passando-se só a considerar os parâmetros condutividade, pH, nitrato, turvação, oxidabilidade, manganês, sódio e potássio.

Tabela 14 – Número de casos em que não foram pedidas análises ou os valores estavam abaixo do limite de quantificação do método utilizado

Parâmetro	Abaixo do Limite de Quantificação	Não Determinado
Condutividade ($\mu\text{S}/\text{cm}$)	0	0
pH (Escala Sørensen)	0	0
Nitrato (mg/dm^3)	10	583
Oxidabilidade (mg/dm^3)	136	580
Manganês (mg/dm^3)	133	587
Turvação (UNT)	149	580
Potássio (mg/dm^3)	0	737
Sódio (mg/dm^3)	0	737
Cobre ($\mu\text{g}/\text{dm}^3$)	21	737
Alumínio ($\mu\text{g}/\text{dm}^3$)	1	779
Ferro ($\mu\text{g}/\text{dm}^3$)	200	580
Nitrito (mg/dm^3)	205	580
Chumbo ($\mu\text{g}/\text{dm}^3$)	50	737
Cádmio ($\mu\text{g}/\text{dm}^3$)	51	737
Níquel ($\mu\text{g}/\text{dm}^3$)	51	737
Amónio (mg/dm^3)	209	580
Crómio ($\mu\text{g}/\text{dm}^3$)	52	737

V.3 – Modelação

Antes de se dar início à fase de construção dos modelos, os objectivos que se pretendem atingir devem estar claramente definidos bem como o conjunto das variáveis de entrada e das variáveis de saída. Como foi referido anteriormente, a frequência de amostragem varia consoante o grupo a que os parâmetros pertencem, sendo maior para o grupo P_1 e menor para o grupo P_3 . Deste modo, tendo em vista caracterizar a água de abastecimento público do Distrito de Évora, procurou-se obter modelos que permitam, com base nos parâmetros químicos do grupo P_1 (pH e condutividade), prever alguns dos parâmetros químicos incluídos no grupo P_2 e P_3 (nitrato, turvação, oxidabilidade, manganês, sódio e potássio).

Atendendo a que se pretende prever valores de variáveis dependentes a partir de exemplos, está-se perante um problema de regressão, tendo-se recorrido às Redes Neurais Artificiais (RNAs). A escolha da referida técnica prende-se com o facto de as RNAs serem, actualmente, consideradas como ferramentas capazes de apreender e representar relações complexas entre variáveis. Para além das razões expostas anteriormente, é ainda de referir que esta técnica tem sido aplicada à problemática relacionada com a qualidade da água, com resultados satisfatórios quer em problemas de classificação [Vicente, 2004; Santos et al., 2005; Atkins et al., 2007; Fernandes, 2009; Fernandes et al., 2009; Pinto et al., 2009; Pinto, 2009; Couto, 2010], quer em problemas de regressão [Karul et al., 2000; Lee et al., 2003; Soyupak et al., 2003; Kuo et al., 2007; Palani et al., 2008; Dogan et al., 2009; Singh et al., 2009].

No que concerne à construção de modelos utilizando RNAs, de entre as várias arquitecturas de rede que se conhecem, optou-se pela utilização de Redes Neurais Unidireccionais (RNUs) uma vez que esta é uma das arquitecturas mais utilizadas tanto em problemas de classificação como de regressão. Como foi referido anteriormente, não existem regras claras para estabelecer quer o número de camadas intermédias, quer o número de nodos que as constituem. Deste modo, no presente trabalho, optou-se pelo procedimento mais utilizado que começa por uma estrutura inicial que vai sendo selectivamente modificada de forma a minimizar uma métrica de erro que, neste trabalho, foi a média do quadrado dos erros [Witten & Frank, 2005].

O algoritmo de treino utilizado foi o algoritmo de retro-propagação que, como foi referido anteriormente, tem sido apontado como o algoritmo de treino supervisionado mais utilizado em RNUs. Pelo mesmo motivo escolheu-se, como função de activação, a função logística (Tabela 6). Em todos os casos optou-se por utilizar os parâmetros de configuração sugeridos pelo *software Weka* (Tabela 15).

No que respeita à implementação de AD foi utilizado o algoritmo C4.5 que, nesta ferramenta, é designado por J.48 tendo-se, tal como anteriormente, utilizado a parametrização sugerida pelo *software* (Tabela 16). Relativamente aos modelos de segmentação, foi utilizada a estratégia *k-means*, sendo apenas necessário fornecer ao programa o número de segmentos a formar.

Tabela 15 – Valores assumidos por defeito pelo algoritmo de implementação de RNAs

Parâmetro	Descrição	Valor por Defeito
<i>Auto Build</i>	Constrói as camadas intermédias da rede	<i>True</i>
<i>Decay</i>	Diminui a taxa de aprendizagem	<i>False</i>
<i>Learning Rate</i>	Taxa de aprendizagem	0,3
<i>Momentum</i>	Coefficiente de momento	0,2
<i>Nominal to Binary Filter</i>	Converte os atributos nominais em binários	<i>True</i>
<i>Normalize Attributes</i>	Normaliza os atributos numéricos	<i>True</i>
<i>Normalize Numeric Classes</i>	Normaliza o atributo a prever caso seja numérico	<i>True</i>
<i>Random Seed</i>	Semente para gerar aleatoriamente os pesos iniciais	0
<i>Reset</i>	Reinicia a aprendizagem, com uma taxa de aprendizagem inferior, caso o algoritmo esteja a divergir	<i>True</i>
<i>Training Time</i>	Número de iterações de treino	500
<i>Validation Set Size</i>	Percentagem de dados a utilizar para validação	0
<i>Validation Threshold</i>	Número de vezes que o erro pode aumentar nos dados de validação até o treino ser concluído	20

Tabela 16 – Valores assumidos por defeito pelo algoritmo de indução de árvores de decisão

Parâmetro	Descrição	Valor por Defeito
<i>Binary Splits</i>	Divisão binária em atributos nominais	<i>False</i>
<i>Confidence Factor</i>	Factor de confiança utilizado na poda	0,25
<i>Debug</i>	Apresentação de informação adicional	<i>False</i>
<i>Min Num Obj</i>	Número mínimo de instâncias por folha	2
<i>Num Folds</i>	Define os dados utilizados para a poda	3
<i>Seed</i>	Semente para gerar aleatoriamente os índices quando se usa a opção <i>reduced error pruning</i>	1
<i>Unpruned</i>	Impede a poda	<i>False</i>

Tendo em vista assegurar a significância estatística dos resultados obtidos, cada simulação foi executada vinte vezes. Em cada uma delas os registos da BD foram aleatoriamente divididos em dois conjuntos mutuamente exclusivos. O conjunto de treino, composto por 67% dos exemplos e o conjunto de teste, constituído pelos restantes.

V.3.1 – Modelos de Segmentação

Tendo em vista procurar tendências, regularidades ou padrões nas análises relativamente a cada uma das variáveis referentes a parâmetros químicos do grupo P₁ (pH e condutividade) utilizaram-se técnicas de aprendizagem não supervisionada, tendo sido construídos modelos de segmentação utilizando a estratégia *k-means*, procurando-se verificar se haveria alguma relação entre a proveniência das amostras e os segmentos formados.

Na Figura 8 mostra-se a distribuição do número de casos pelos vários concelhos do Distrito de Évora. A sua análise mostra que os concelhos de Évora, Estremoz, Arraiolos e Reguengos de Monsaraz são os que estão mais representados na BD. Os concelhos de Mourão e Mora, por seu lado, são os menos representados.

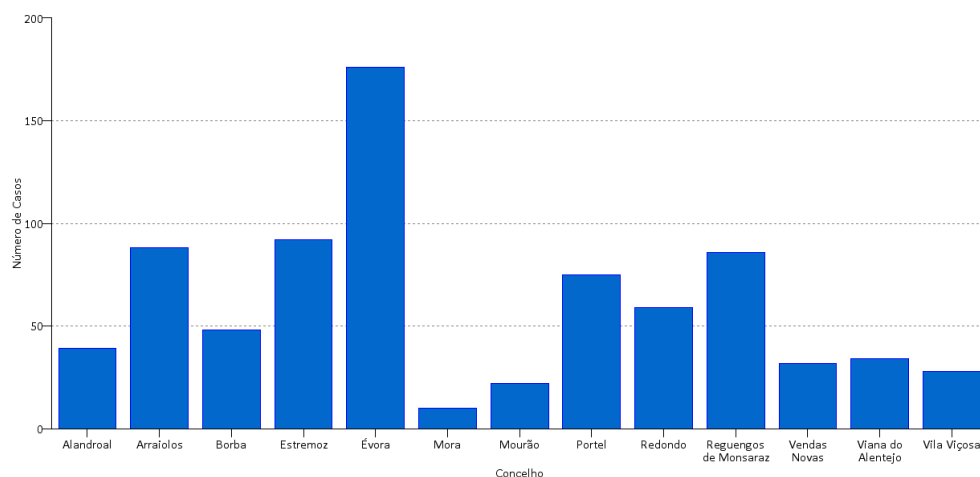


Figura 8 – Distribuição do número de casos pelos concelhos do Distrito de Évora

Na abordagem *k-means*, como foi referido anteriormente, é necessário fornecer o parâmetro k , que corresponde ao número de segmentos a formar, tendo-se utilizado valores entre dois e quatro. Quando se utilizaram valores de k superiores, alguns segmentos eram constituídos por um número muito reduzido de casos, não permitindo a sua análise posterior. Por este motivo, optou-se por não os apresentar. Na Tabela 17 apresenta-se o número de casos para cada modelo ensaiado.

Tabela 17 – Número de objectos por segmento para os vários modelos estudados

Modelo		Número de Objectos
$k = 2$	Segmento 1	365
	Segmento 2	424
$k = 3$	Segmento 1	358
	Segmento 2	90
	Segmento 3	341
$k = 4$	Segmento 1	286
	Segmento 2	90
	Segmento 3	102
	Segmento 4	311

Tendo em vista perceber como é que foram atribuídos os casos aos vários segmentos geraram-se modelos explicativos da segmentação, tendo-se optado por utilizar Árvores de Decisão (AD). Os motivos subjacentes a essa escolha prendem-se com o seu elevado poder expressivo e com a facilidade de interpretação e de utilização.

A aplicação da estratégia *k-means*, com $k = 2$, aos parâmetros físico-químicos anteriormente referidos originou dois segmentos com, respectivamente, 365 e 424 registos. Os centros de gravidade de cada um dos segmentos encontram-se na Tabela 18. As variáveis cujo índice de importância é próximo da unidade encontram-se destacadas a negrito. O referido índice é calculado como a diferença entre a unidade e o valor de significância do teste *t-Student*. A hipótese nula considera que as médias das variáveis são as mesmas para todos os segmentos. Deste modo, se a variável influenciar a definição do segmento a hipótese nula é rejeitada, o valor da significância do teste *t-Student* é próximo de zero e a importância é próxima da unidade.

A análise da Tabela 18 mostra que o segmento 1 concentra as amostras cuja condutividade é mais elevada e portanto poderá estar relacionada com a zona onde a amostra foi colhida. Na perspectiva de testar a hipótese anteriormente formulada construiu-se o gráfico apresentado na Figura 9 onde se mostra, em termos percentuais, as associações entre a proveniência das amostras e os segmentos formados. A sua análise mostra que o segmento 1 contém a totalidade das amostras provenientes do concelho de Borba, a maior parte das amostras referentes aos concelhos de Arraiolos, Estremoz, Mourão, Viana do Alentejo e Alandroal. O segmento 2 é formado por todas as amostras provenientes do concelho de Vendas Novas, pela maioria das amostras provenientes dos concelhos de Reguengos de Monsaraz, Mora, Redondo e Évora. A análise do gráfico permite salientar, ainda, que as amostras dos concelhos de Vila Viçosa e Portel se dividem por ambos os segmentos.

Tabela 18 – Centros de gravidade dos segmentos obtidos com $k = 2$

Variável	Segmento 1	Segmento 2
Condutividade ($\mu\text{S}/\text{cm}$)	$699,7 \pm 115,2$	$387,8 \pm 85,8$
pH (Escala Sørensen)	$7,6 \pm 0,2$	$7,5 \pm 0,4$

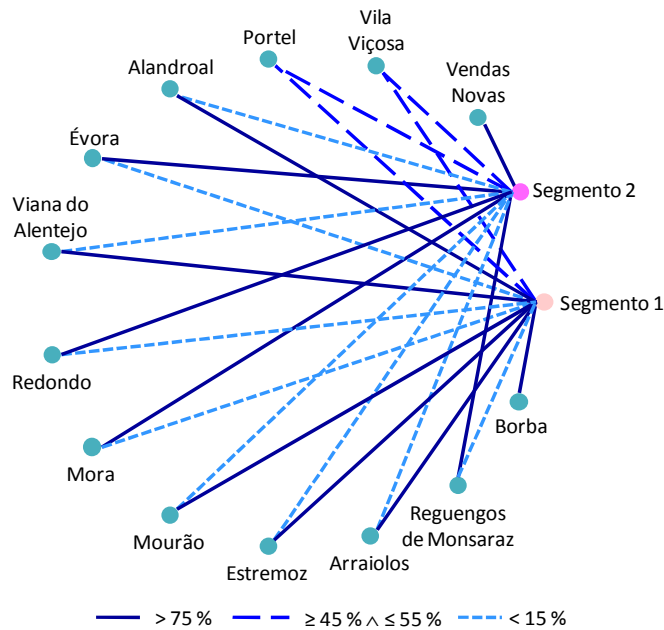


Figura 9 – Associação entre os segmentos e os concelhos de onde são provenientes as amostras

A AD explicativa do modelo de segmentação encontra-se na Figura 10 onde se pode inferir que, no segmento 1, estão as amostras cuja condutividade é superior a 540 $\mu\text{S}/\text{cm}$. A acuidade do modelo, medida em termos de percentagem de acertos relativamente ao número de casos apresentados, é de 99,4 % (533 casos correctamente classificados em 536) para o conjunto de treino e de 100 % para o conjunto de teste, Tabela 19.

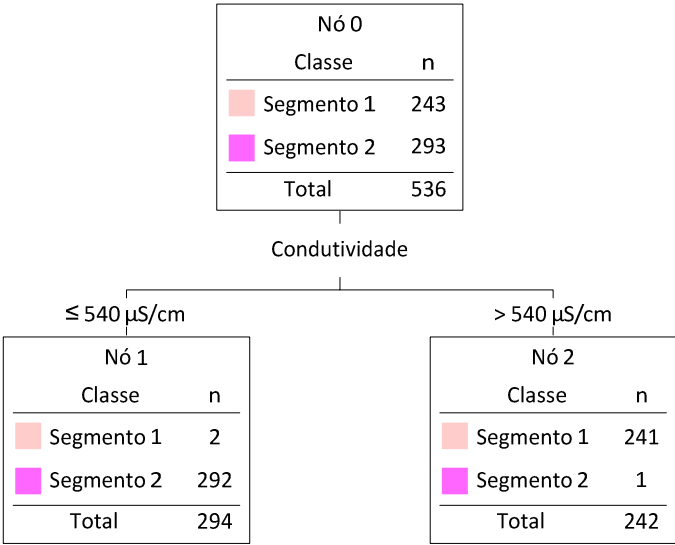


Figura 10 – Árvore de decisão explicativa do modelo de segmentação ($k = 2$) com base nos valores da condutividade e do pH

Tabela 19 – Matriz de coincidências relativa ao modelo apresentado na Figura 10

Segmento	Conjunto de Treino		Conjunto de Teste	
	1	2	1	2
1	241	2	122	0
2	1	292	0	131

A aplicação da estratégia *k-means*, com $k = 3$, aos parâmetros físico-químicos anteriormente referidos originou três segmentos com, respectivamente, 358, 90 e 341 registos. Os centros de gravidade de cada um dos segmentos encontram-se na Tabela 20. As variáveis cujo índice de importância é próximo de um encontram-se destacadas a negrito. A sua análise mostra que o segmento 3 concentra as amostras cuja condutividade é mais elevada, estando as amostras cuja condutividade é mais baixa no segmento 2.

Para verificar se os segmentos estão relacionados com a zona onde a amostra foi colhida construiu-se o gráfico apresentado na Figura 11 onde se mostra, em termos percentuais, as associações entre a proveniência das amostras e os segmentos formados. A sua análise mostra que o segmento 1 contém a maior parte das amostras referentes aos concelhos de Redondo, Reguengos de Monsaraz e Évora. O segmento 2 é formado pela maioria das amostras provenientes do concelho de Vendas Novas e, finalmente, o segmento 3 é formado pela totalidade das amostras oriundas do concelho de Borba, e pela maioria das amostras provenientes dos concelhos de Mourão, Arraiolos e Estremoz. A análise do gráfico permite ainda salientar que as amostras do concelho do Alandroal e a maioria das amostras dos concelhos de Vila Viçosa, Viana do Alentejo e Portel se dividem pelos segmentos 1 e 3 enquanto as amostras colhidas no concelho de Mora se distribuem maioritariamente pelos segmentos 1 e 2.

Tabela 20 – Centros de gravidade dos segmentos obtidos com $k = 3$

Variável	Segmento 1	Segmento 2	Segmento 3
Condutividade ($\mu\text{S}/\text{cm}$)	422,1 \pm 66,7	297,1 \pm 109,9	709,6 \pm 112,8
pH (Escala Sørensen)	7,7 \pm 0,3	6,9 \pm 0,4	7,5 \pm 0,2

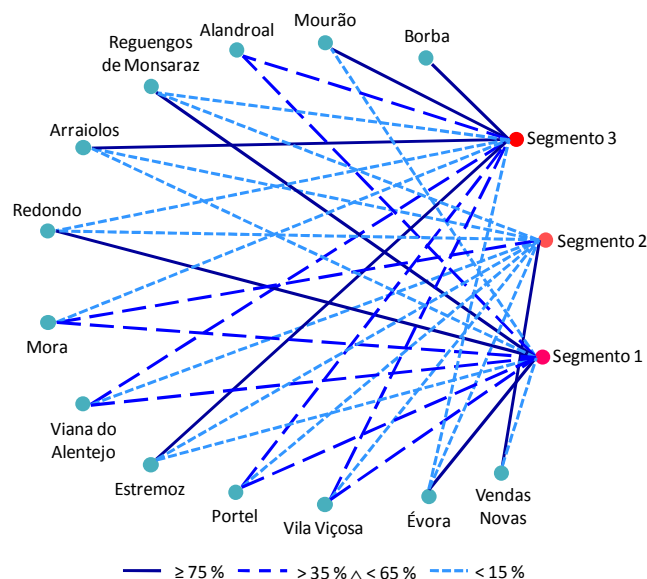


Figura 11 – Associação entre os segmentos e os concelhos de onde são provenientes as amostras

A AD explicativa deste modelo de segmentação encontra-se na Figura 12. A acuidade do modelo, medida em termos de percentagem de acertos relativamente ao número de casos apresentados, é de 99,6 % (534 casos correctamente classificados em 536) para o conjunto de treino e de 100 % para o conjunto de teste, Tabela 21.

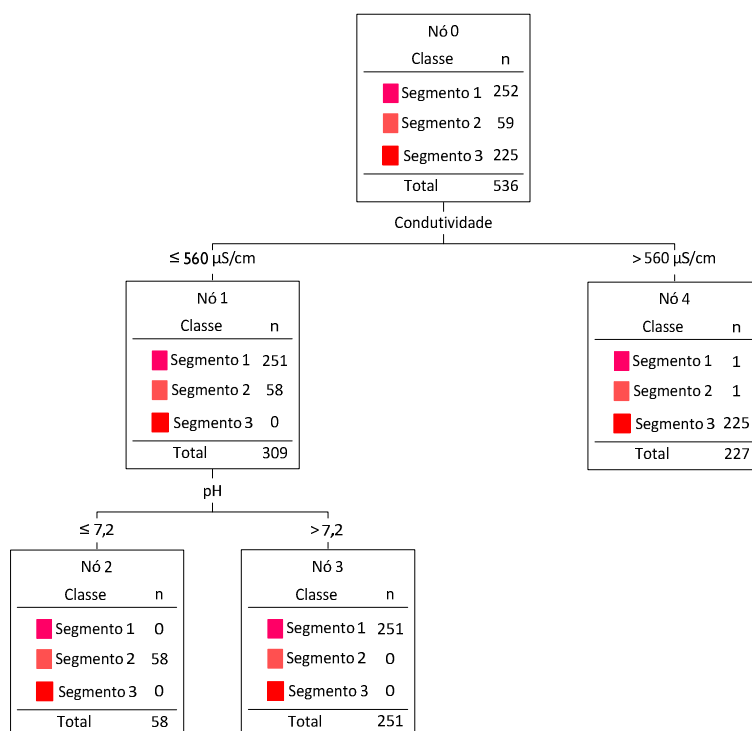


Figura 12 – Árvore de decisão explicativa do modelo de segmentação ($k = 3$) com base nos valores da condutividade e do pH

Tabela 21 – Matriz de coincidências relativa ao modelo apresentado na Figura 12

Segmento	Conjunto de Treino			Conjunto de Teste		
	1	2	3	1	2	3
1	251	0	1	106	0	0
2	0	58	1	0	31	0
3	0	0	225	0	0	116

A aplicação da estratégia *k-means*, com $k = 4$, aos parâmetros físico-químicos anteriormente referidos originou quatro segmentos com, respectivamente, 286, 90, 102 e 311 registos. Os centros de gravidade dos segmentos encontram-se na Tabela 22, onde se destacam, a negrito, as variáveis cujo índice de importância é próximo da unidade.

A análise da Tabela 22 mostra que o segmento 3 concentra as amostras cuja condutividade é mais elevada, estando as amostras cuja condutividade é mais baixa no segmento 2. Para verificar se os segmentos estão relacionados com a zona onde a amostra foi colhida construiu-se o gráfico apresentado na Figura 13 onde se mostra, em termos percentuais, as associações entre a proveniência das amostras e os segmentos formados.

A sua análise mostra que o segmento 1 contém a totalidade das amostras do concelho de Borba, a maior parte das amostras dos concelhos de Vila Viçosa, Alandroal, Estremoz e Viana do Alentejo. O segmento 2, tal como anteriormente, é formado pela maioria das amostras provenientes do concelho de Vendas Novas. O segmento 4 é formado pela maioria das amostras provenientes dos concelhos de Redondo, Évora e Reguengos de Monsaraz. A análise do gráfico permite salientar, ainda que as amostras dos concelhos de Arraiolos e Mourão se dividem maioritariamente pelos segmentos 1 e 3 enquanto a maioria das amostras colhidas no concelho de Mora se distribuem pelos segmentos 2 e 4. As amostras referentes ao concelho de Portel distribuem-se maioritariamente pelos segmentos 1, 3 e 4.

Tabela 22 – Centros de gravidade dos segmentos obtidos com $k = 4$

Variável	Segmento 1	Segmento 2	Segmento 3	Segmento 4
Condutividade ($\mu\text{S}/\text{cm}$)	631,5 \pm 60,7	297,1 \pm 109,9	849,8 \pm 95,6	404,5 \pm 51,9
pH (Escala Sørensen)	7,6 \pm 0,2	6,9 \pm 0,4	7,5 \pm 0,2	7,7 \pm 0,3

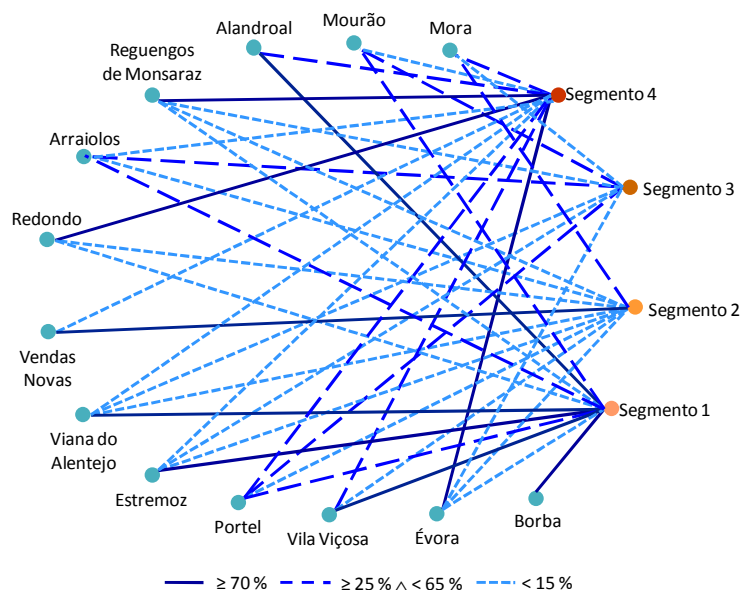


Figura 13 – Associação entre os segmentos e os concelhos de onde são provenientes as amostras

A AD explicativa deste modelo de segmentação encontra-se na Figura 14. A acuidade do modelo, medida em termos de percentagem de acertos relativamente ao número de casos apresentados, é de 99,4 % (533 casos correctamente classificados em 536) para o conjunto de treino e de 99,6 % (252 casos correctamente classificados em 253) para o conjunto de teste, Tabela 23.

Uma análise conjunta dos resultados apresentados anteriormente, relativos aos três modelos de segmentação, permite destacar semelhanças e diferenças que possibilitam discutir os méritos dos diferentes modelos. Na Figura 15 mostra-se, de uma forma esquemática, o modo como os segmentos dos vários modelos se articulam. Apresenta-se, ainda, as origens das amostras destacando-se, a negrito, os concelhos com mais de 70 % de amostras num dado segmento, omitindo-se os concelhos que contribuem com menos de 15 % de amostras.

Tendo em consideração a caracterização das origens da água de abastecimento público do Distrito de Évora, que se apresenta no Capítulo I, e a análise da Figura 15 verifica-se que, no modelo de segmentação obtido com $k = 2$, o segmento 1 integra amostras provenientes de concelhos abastecidos por águas subterrâneas enquanto o segmento 2, por seu lado, é constituído por amostras dos concelhos abastecidos por águas superficiais com excepção dos concelhos de Mora e de Vendas Novas.

Quando se passa para o modelo de três segmentos, o novo segmento integra amostras provenientes do concelho de Vendas Novas e parte das oriundas do concelho de Mora caracterizadas por apresentar os valores de condutividade e de pH mais baixos. Deste modo, o modelo de três segmentos permite separar as referidas amostras, de origem subterrânea, das amostras de origem superficial. O aumento da complexidade do modelo, com a introdução de um quarto segmento, não parece trazer benefícios ao nível da interpretação dos resultados. Contudo, são de salientar dois aspectos. O primeiro prende-se com o facto do segmento 2 se manter inalterado e, o segundo, está relacionado com o facto de o novo segmento integrar as amostras de origem subterrânea cuja condutividade é mais elevada, superior a $730 \mu\text{S}/\text{cm}$.

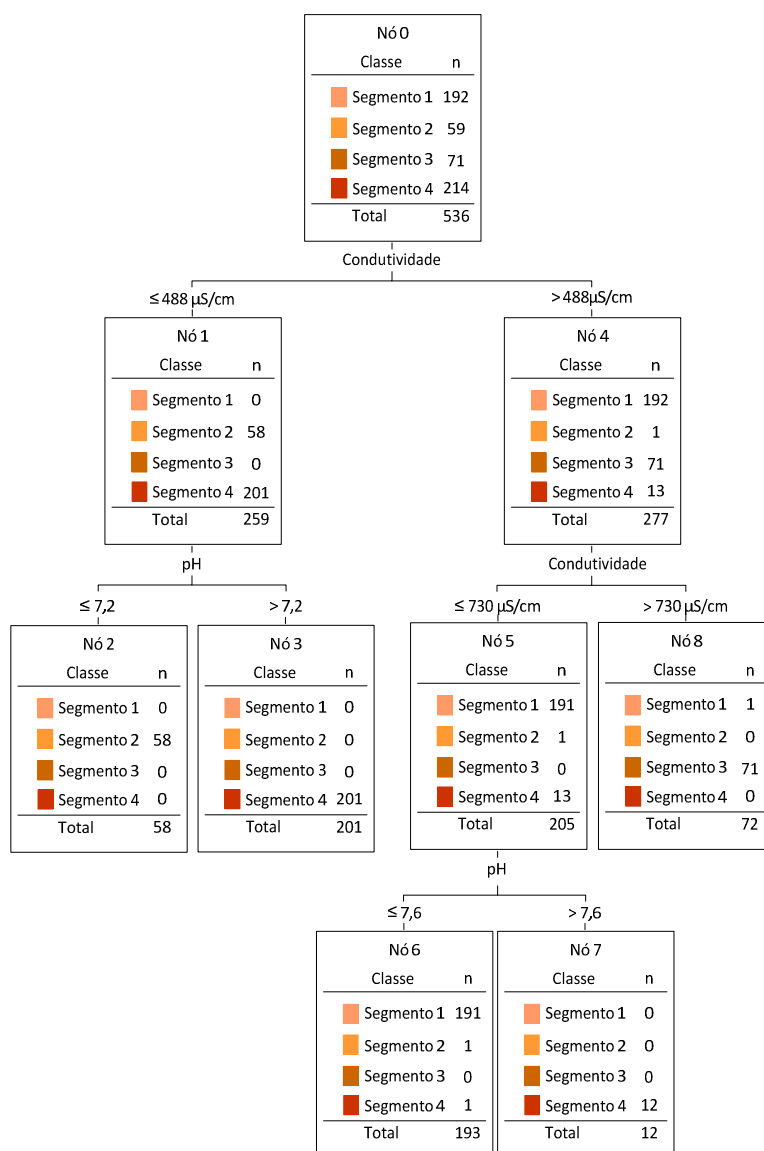


Figura 14 – Árvore de decisão explicativa do modelo de segmentação ($k = 4$) com base nos valores da condutividade e do pH

Tabela 23 – Matriz de coincidências relativa ao modelo apresentado na Figura 14

	Conjunto de Treino				Conjunto de Teste			
Segmento	1	2	3	4	1	2	3	4
1	191	0	1	0	94	0	0	0
2	1	58	0	0	0	30	0	1
3	0	0	71	0	0	0	31	0
4	1	0	0	213	0	0	0	97

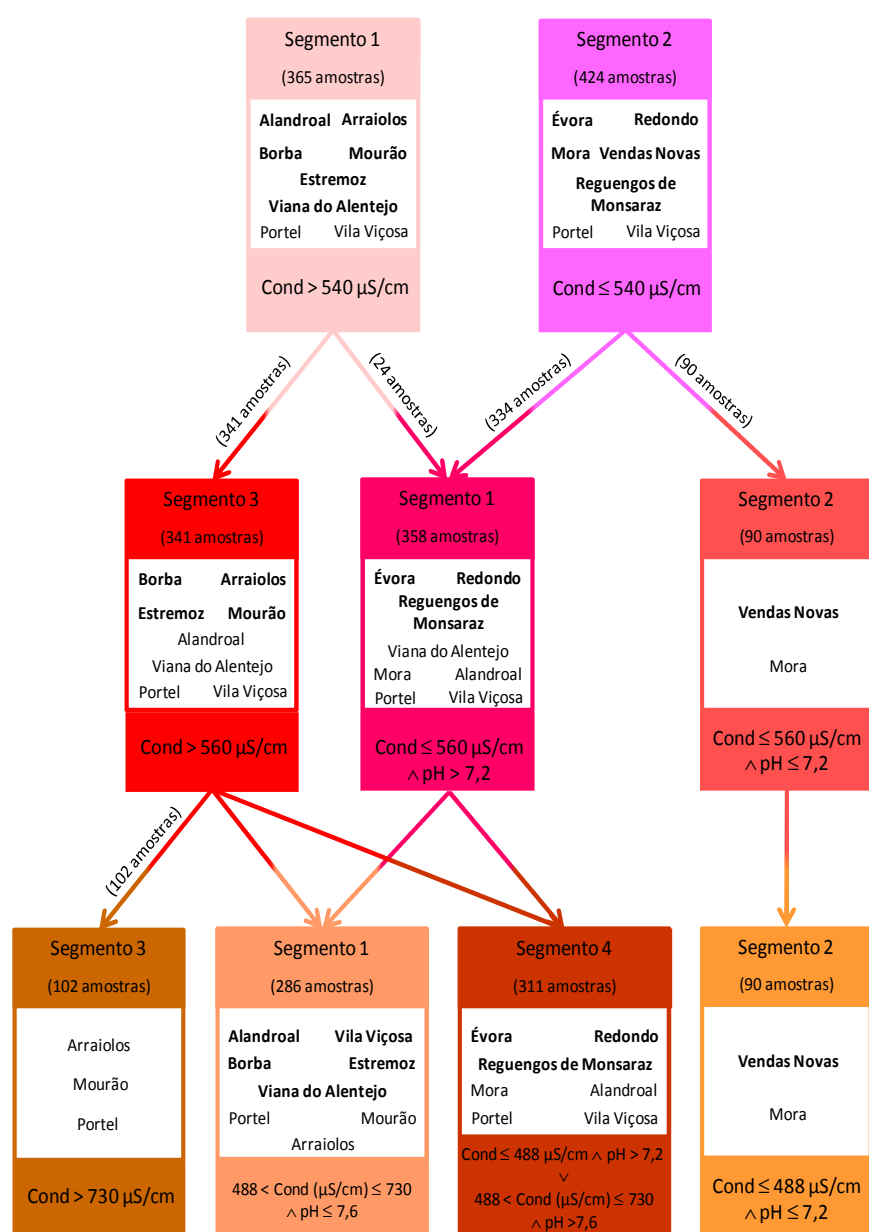


Figura 15 – Articulação entre os segmentos dos diversos modelos de segmentação apresentados

Atendendo ao que acabou de ser exposto, o modelo de três segmentos parece ser o mais adequado, uma vez que é relativamente simples e consegue agrupar as amostras provenientes de águas subterrâneas no segmento 3, as amostras originárias de águas superficiais no segmento 1 e, no segmento 2, amostras de água subterrânea que do ponto de vista da condutividade se apresentam próximas das águas de origem superficial distinguindo-se, contudo, pelos valores de pH mais baixos. Neste modelo de três segmentos as amostras oriundas dos concelhos de Alandroal, Portel, Viana do Alentejo e Vila Viçosa surgem divididas pelos segmentos 1 e 3. Na verdade, no caso do concelho de Portel a água dos fontanários públicos, como foi referido no capítulo I, tem origem subterrânea e a água da rede é produzida a partir de águas superficiais. Relativamente aos concelhos de Viana do Alentejo e do Alandroal a água utilizada na produção de água para consumo humano resulta da mistura de águas subterrâneas e superficiais. O caso do concelho de Vila Viçosa é um pouco diferente uma vez que a água da rede resulta da mistura de água subterrânea do aquífero Estremoz – Cano com água de outros furos que, possivelmente, alteram as suas características. Estes factos podem ser tomados como pontos fortes do modelo uma vez que realçam a sua capacidade de reconhecer e separar as amostras em função da origem, apenas com base na condutividade e no pH.

V.3.2 – Modelos de Regressão

No que concerne ao desenvolvimento de modelos de regressão verificou-se que, mesmo só considerando os parâmetros nitrato, turvação, oxidabilidade, manganês, potássio e sódio, o número de registos com informação completa era escasso, não permitindo desenvolver modelos capazes de prever simultaneamente todos eles. Assim, optou-se por desenvolver modelos com diferentes variáveis dependentes, sendo estas determinadas pelo número de registos em que havia informação simultânea sobre pelo menos duas. Na Tabela 24 apresenta-se o número de registos em que existia informação completa relativa a cada par de parâmetros. A sua análise mostra que, quase sempre, há registos com informação relativa a nitrato e a qualquer um dos outros parâmetros. Por exemplo, existem 60 registos com informação relativa à turvação (Tabela 13) e em 52 deles (86,7%) existe, também, informação relativa a nitrato. Deste modo, desenvolveram-se quatro tipos de modelos nos quais as variáveis dependentes são nitrato, sódio e potássio; nitrato e oxidabilidade; nitrato e manganês e, finalmente, nitrato e turvação.

Para o efeito construíram-se quatro BD com os dados referentes a cada um dos tipos de modelos a desenvolver. Na Tabela 25 apresenta-se a caracterização estatística das variáveis presentes em cada BD.

Tabela 24 – Número de registos com informação completa relativa a cada par de parâmetros

	Turvação	Oxidabilidade	Manganês	Sódio	Potássio
Nitrato	52	62	60	52	52
Turvação	---	29	32	12	12
Oxidabilidade	---	---	46	11	11
Manganês	---	---	---	14	14
Sódio	---	---	---	---	52

Tabela 25 – Caracterização estatística das variáveis presentes em cada uma das bases de dados

	Variável	Mínimo	Máximo	Média	Desvio Padrão
Base de Dados 1 (52 registos)	pH (Escala Sørensen)	6,3	8,4	7,5	0,3
	Condutividade ($\mu\text{S}/\text{cm}$)	152,0	900,0	563,2	180,3
	Nitrato (mg/dm^3)	1,7	51,0	19,7	14,5
	Potássio (mg/dm^3)	0,4	9,0	2,6	1,5
	Sódio (mg/dm^3)	9,0	240,0	37,2	35,0
Base de Dados 2 (62 registos)	pH (Escala Sørensen)	7,2	8,7	7,7	0,3
	Condutividade ($\mu\text{S}/\text{cm}$)	322,0	910,0	422,5	117,1
	Nitrato (mg/dm^3)	1,7	49,0	8,1	9,6
	Oxidabilidade (mg/dm^3)	0,9	3,8	2,0	0,6
Base de Dados 3 (60 registos)	pH (Escala Sørensen)	6,7	8,7	7,7	0,4
	Condutividade ($\mu\text{S}/\text{cm}$)	255,0	1220,0	478,5	186,5
	Nitrato (mg/dm^3)	1,6	130,0	12,9	21,8
	Manganês (mg/dm^3)	2,0	48,0	11,8	8,1

Tabela 25 (Cont.) – Caracterização estatística das variáveis presentes em cada uma das bases de dados

	Variável	Mínimo	Máximo	Média	Desvio Padrão
Base de Dados 4 (52 registos)	pH (Escala Sørensen)	7,2	8,1	7,6	0,3
	Condutividade (μS/cm)	304,0	960,0	505,8	162,3
	Nitrato (mg/dm ³)	2,1	78,0	14,1	15,7
	Turvação (UNT)	0,2	5,1	0,8	1,1

Outro aspecto importante, que deve ser atendido durante a fase de pré-processamento, prende-se com o escalonamento dos dados. Este procedimento visa transformá-los de modo a acelerar e melhorar o processo de aprendizagem. Na verdade, os algoritmos de aprendizagem de gradiente descendente, como o algoritmo de retro-propagação, são bastante sensíveis ao escalonamento das variáveis. Por exemplo, se uma saída tem valores entre 0,001 e 0,1 e uma outra tem valores entre 0 e 5000, o algoritmo de treino vai dar mais importância à aprendizagem da segunda saída [Pyle, 1999; Han & Kamber, 2006; Bramer, 2007]. Neste trabalho os dados foram escalonados para o intervalo [0, 1] tendo sido utilizada a equação:

$$X_{esc} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (14)$$

onde X , X_{esc} , X_{max} e X_{min} denotam, respectivamente, o valor de atributo, o valor de atributo escalonado, o valor máximo e o valor mínimo que o atributo apresenta.

V.3.2.1 – Modelos de previsão dos parâmetros nitrato, potássio e sódio

Para a modelação dos parâmetros nitrato, sódio e potássio, com base nos valores de pH e da condutividade elaboraram-se e testaram-se algumas topologias de rede. Na Tabela 26 encontram-se os valores do Desvio Absoluto Médio (DAM), da Média do Quadrado dos Erros (MQE) e de *bias* para algumas das topologias testadas, tanto para o conjunto de treino como para o conjunto de teste.

Tabela 26 – Valores do DAM, da MQE e de *bias* para algumas das topologias testadas

Topologia	Nitrato		Potássio		Sódio	
	Treino	Teste	Treino	Teste	Treino	Teste
DESVIO ABSOLUTO MÉDIO (DAM)						
2-17-13-3	1,486	1,334	0,174	0,176	6,618	3,997
2-17-14-3	1,477	1,523	0,248	0,153	7,236	7,321
2-18-14-3	0,503	0,381	0,063	0,056	1,701	2,297
2-18-15-3	1,417	0,927	0,173	0,195	5,407	5,057
2-19-17-3	1,551	0,670	0,153	0,159	3,863	3,580
2-22-18-3	1,289	0,802	0,159	0,260	7,098	5,612
2-22-22-3	1,524	0,869	0,252	0,105	5,489	5,216
MÉDIA DO QUADRADO DOS ERROS (MQE)						
2-17-13-3	4,135	3,961	0,060	0,112	67,531	29,946
2-17-14-3	4,555	3,967	0,109	0,045	92,179	75,065
2-18-14-3	0,380	0,231	0,005	0,004	4,043	6,530
2-18-15-3	3,866	2,148	0,052	0,088	56,764	35,668
2-19-17-3	5,352	1,469	0,042	0,100	28,000	19,296
2-22-18-3	3,768	1,049	0,046	0,156	84,126	48,043
2-22-22-3	4,527	1,388	0,127	0,022	65,793	47,394
BIAS						
2-17-13-3	-0,072	0,369	0,013	-0,098	-0,513	0,015
2-17-14-3	-0,338	0,378	0,042	-0,038	-0,730	1,330
2-18-14-3	-0,096	0,047	0,010	-0,003	0,155	0,158
2-18-15-3	-0,205	0,295	0,029	-0,004	0,036	0,843
2-19-17-3	-0,244	0,188	0,004	-0,080	0,228	-1,075
2-22-18-3	-0,308	0,301	0,026	-0,076	0,121	0,653
2-22-22-3	-0,464	0,320	0,025	0,012	-0,116	0,080

A análise da Tabela 26 mostra que a topologia 2-18-14-3 é aquela que minimiza tanto o DAM como a MQE revelando-se, de entre as topologias testadas, aquela que apresenta melhor desempenho na previsão dos teores em nitrato, sódio e potássio. Relativamente ao valor de *bias* (erro médio) pode afirmar-se que, para a referida topologia, está próximo de zero não apresentando, por isso, tendência para subestimar ou para sobrestimar as variáveis de saída.

A RNA seleccionada para modelar o teor de nitrato, potássio e sódio nos sistemas de abastecimento de água para consumo humano do Distrito de Évora, a partir dos valores do pH e de condutividade é apresentada na Figura 16. Consiste numa camada de entrada com dois nós, correspondentes às variáveis condutividade e pH, duas camadas intermédias com, respectivamente, dezoito e catorze nós e, por último, uma camada de saída com três nós, correspondentes às variáveis nitrato, potássio e sódio.

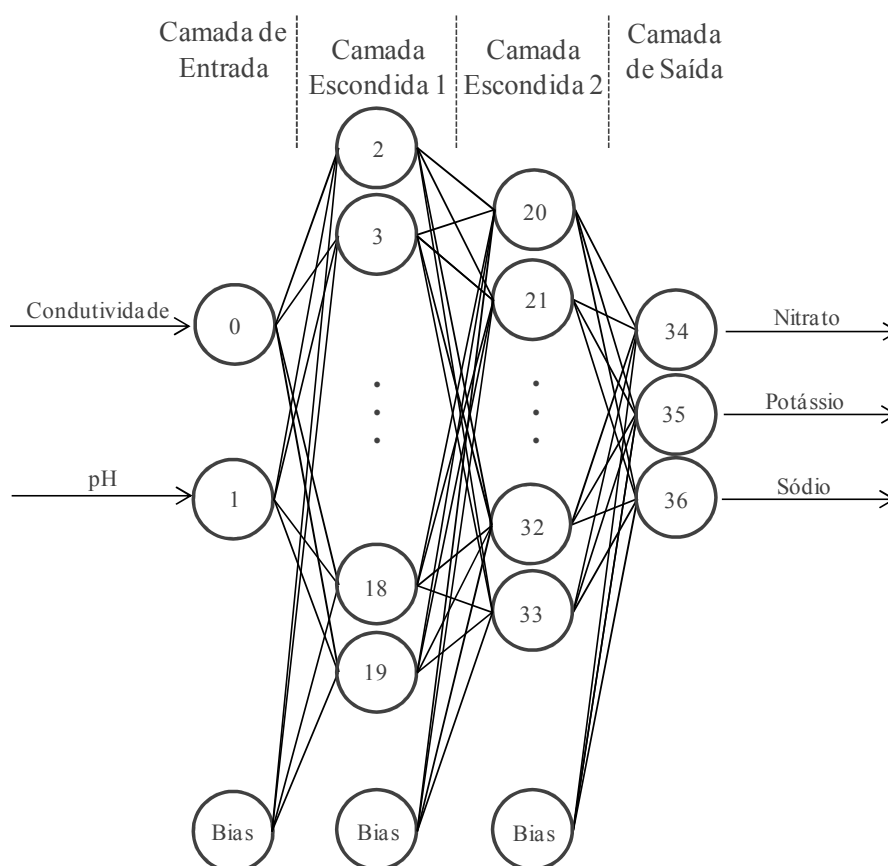


Figura 16 – RNA seleccionada para a previsão dos parâmetros nitrato, potássio e sódio a partir dos parâmetros pH e condutividade

No que concerne à avaliação do desempenho do modelo seleccionado apresenta-se a representação gráfica dos valores por ele previstos em função dos valores experimentais, tanto para o conjunto de treino como para o conjunto de teste (Figura 17) indicando-se, para cada uma delas, o valor do coeficiente de determinação (R^2). Quer a análise dos gráficos apresentados na Figura 17 quer os valores de R^2 obtidos, que variam entre 0,9981 e 0,9905 para o conjunto de treino e entre 0,9976 e 0,9991 para o conjunto de teste, atestam a capacidade do modelo na previsão dos valores de nitrato, potássio e de sódio, a partir do conhecimento da condutividade e do pH, em todo o intervalo de valores ensaiado. Um outro aspecto a salientar baseia-se no facto de os valores de R^2 relativos aos conjuntos de treino e de teste serem semelhantes. Este resultado parece mostrar que o modelo não apresenta problemas de sobre-ajustamento.

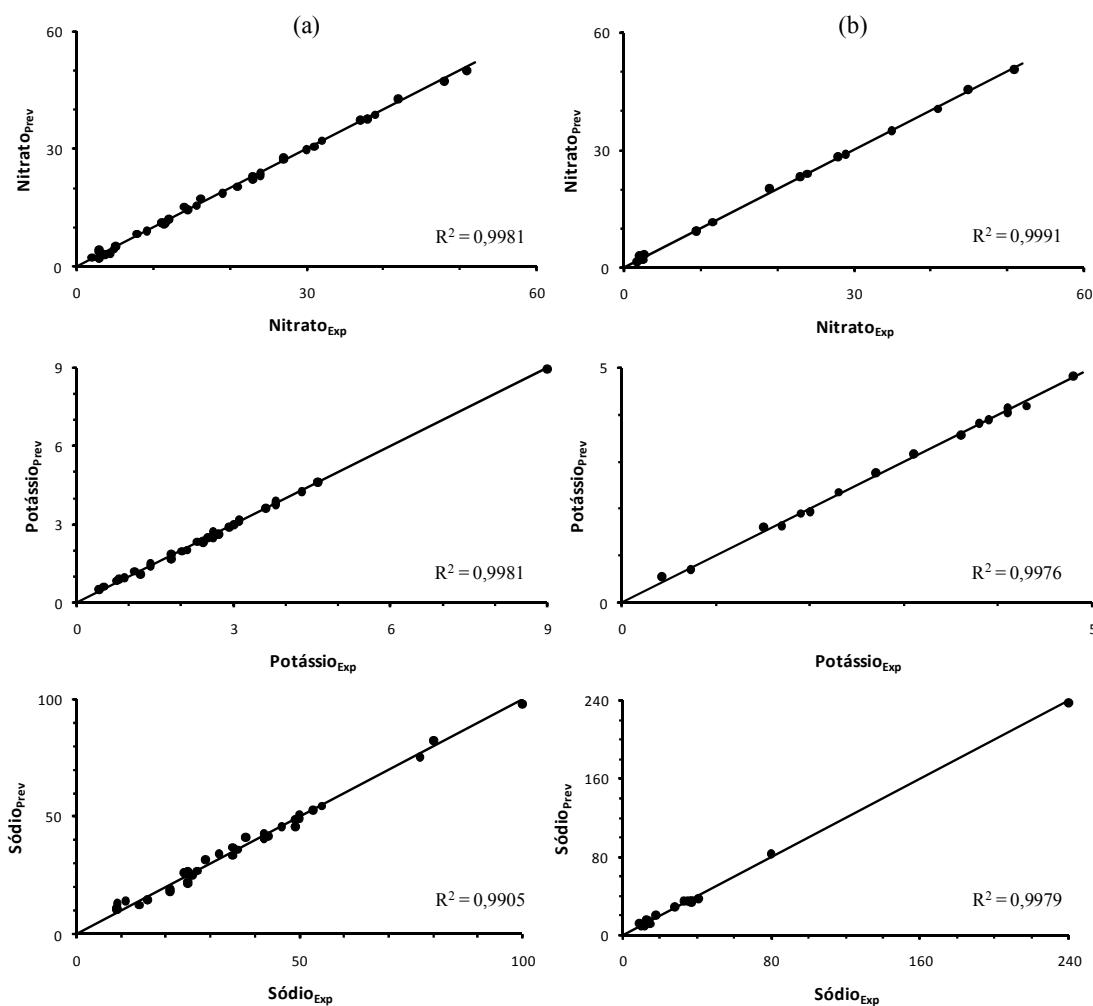


Figura 17 – Valores previstos vs valores experimentais das variáveis de saída para o conjunto de treino (a) e para o conjunto de teste (b)

Como foi anteriormente destacado, os valores de *bias* parecem mostrar que o modelo seleccionado não apresenta tendência para subestimar ou sobrestimar as variáveis de saída. Para confirmar tal facto construíram-se os gráficos que se apresentam na Figura 18, onde se representou o erro ou resíduo em funções dos valores previstos pelo modelo, para cada uma das variáveis de saída, tanto para o conjunto de treino como para o conjunto de teste.

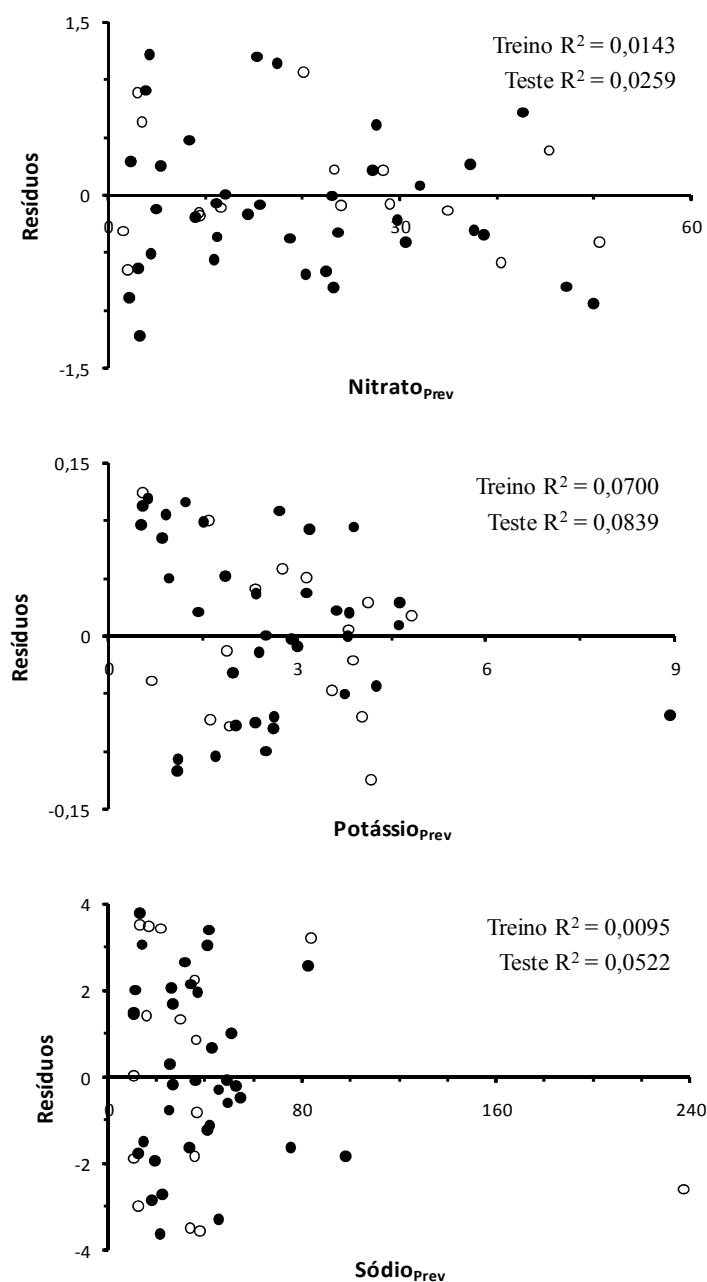


Figura 18 – Resíduos vs valores previstos para as variáveis de saída.

(●) Conjunto de treino (○) conjunto de teste

A análise quer dos gráficos da Figura 18 quer dos valores de R^2 , que nela se apresentam, mostra que os resíduos se distribuem de modo aleatório, de ambos os lados da linha horizontal e são independentes do valor da variável de saída. Estes resultados corroboram a ideia, apresentada anteriormente, que o modelo tem elevado desempenho na previsão dos valores de nitrato, potássio e sódio, a partir do conhecimento da condutividade e do pH, em todo o intervalo de valores ensaiado. Na verdade, se a distribuição dos resíduos não fosse aleatória e independente do valor da variável de saída significaria que o modelo não se ajustava adequadamente aos dados [Singh et al., 2009; NIST/SEMATECH, 2010].

O número de registos existentes na BD original com informação completa relativa aos parâmetros nitrato, potássio e sódio não permite ter um conjunto independente de dados para validação. Contudo, atendendo ao facto de, por um lado, haver 196 registos com informação completa relativa ao parâmetro nitrato (Tabela 13) e, por outro, apenas 52 desses registos serem utilizados para treinar e testar o modelo anteriormente apresentado (Tabela 24), utilizaram-se os 144 casos restantes para o validar relativamente à sua capacidade para prever os valores da concentração de nitrato a partir dos valores de pH e de condutividade. Para tal forneceram-se os valores de pH e de condutividade e compararam-se os valores previstos com os valores experimentais (Figura 19). A sua análise bem como do valor de R^2 , que nela se apresenta, mostra que, mesmo para dados independentes, o desempenho do modelo se mantém elevado no que respeita à previsão do parâmetro nitrato.

A distribuição dos resíduos em função dos valores de nitrato previstos, para os dados de validação, encontra-se na Figura 20 verificando-se que, tal como anteriormente, são independentes do valor da saída e se distribuem aleatoriamente. Estes resultados mostram que o modelo não apresenta problemas de sobre-ajustamento i.e. ter um desempenho quase perfeito nos dados de treino, mas um desempenho pobre em dados independentes.

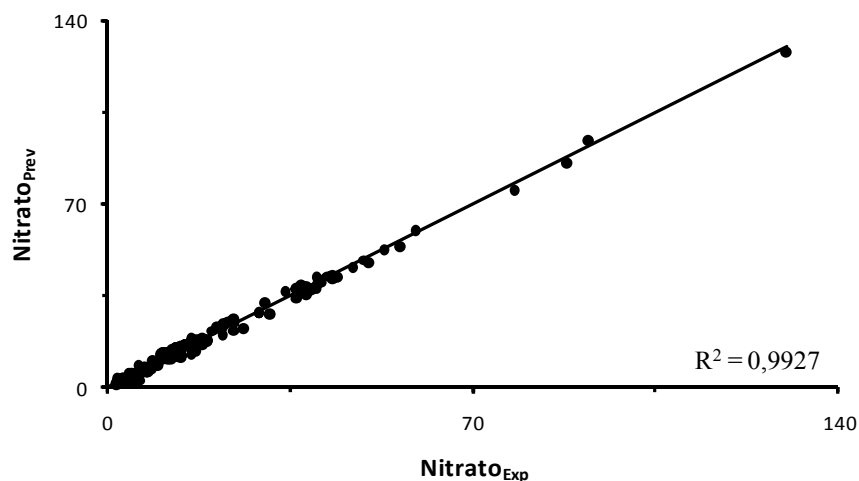


Figura 19 – Valores previstos vs valores experimentais relativos ao conjunto de validação

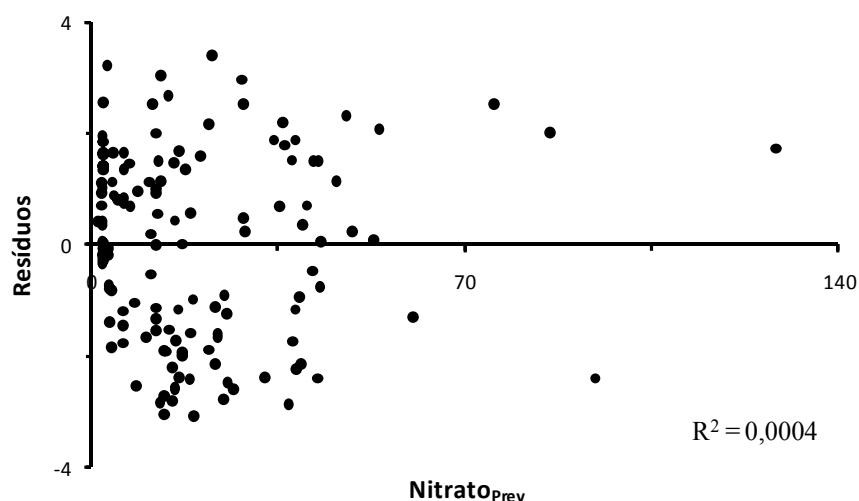


Figura 20 – Resíduos vs valores previstos relativos ao conjunto de validação

V.3.2.2 – Modelos de previsão dos parâmetros nitrato e oxidabilidade

Para a previsão dos parâmetros nitrato e oxidabilidade, utilizando os valores de pH e da condutividade, elaboraram-se e testaram-se algumas topologias de rede cujos valores do DAM, da MQE e de *bias* se apresentam na Tabela 27, tanto para o conjunto de treino como para o conjunto de teste.

Tabela 27 – Valores do DAM, da MQE e de *bias* para algumas das topologias testadas

Topologia	Nitrato		Oxidabilidade	
	Treino	Teste	Treino	Teste
DESVIO ABSOLUTO MÉDIO (DAM)				
2-12-9-2	0,982	0,840	0,062	0,049
2-13-11-2	0,785	1,113	0,050	0,057
2-13-12-2	0,870	0,899	0,073	0,065
2-16-14-2	1,174	0,749	0,056	0,115
2-19-10-2	0,371	0,334	0,049	0,045
2-20-12-2	0,687	0,612	0,026	0,019
2-21-13-2	0,909	0,684	0,069	0,130
2-24-12-2	1,148	1,314	0,087	0,076
MÉDIA DO QUADRADO DOS ERROS (MQE)				
2-12-9-2	2,558	1,201	0,007	0,006
2-13-11-2	2,710	3,850	0,005	0,006
2-13-12-2	1,972	1,920	0,011	0,007
2-16-14-2	3,163	1,082	0,011	0,114
2-19-10-2	0,194	0,178	0,005	0,005
2-20-12-2	0,784	0,520	0,001	0,001
2-21-13-2	2,566	0,908	0,012	0,116
2-24-12-2	3,738	3,436	0,024	0,014

Tabela 27 (Cont.) – Valores do DAM, da MQE e de bias para algumas das topologias testadas

Topologia	Nitrato		Oxidabilidade	
	Treino	Teste	Treino	Teste
<i>BIAS</i>				
2-12-9-2	-0,177	0,115	-0,008	0,010
2-13-11-2	-0,024	-0,290	-0,006	0,015
2-13-12-2	-0,224	-0,058	-0,010	0,013
2-16-14-2	-0,461	0,285	-0,014	-0,068
2-19-10-2	-0,005	-0,058	-0,010	-0,007
2-20-12-2	-0,049	0,181	-0,005	0,004
2-21-13-2	-0,131	0,237	-0,019	-0,076
2-24-12-2	-0,372	-0,259	-0,022	0,040

Através da análise da Tabela 27 verifica-se que a topologia 2-19-10-2 é aquela que minimiza tanto o DAM como a MQE revelando-se, de entre as topologias testadas, aquela que apresenta melhor desempenho na previsão dos teores em nitrato. No entanto, relativamente à previsão do parâmetro oxidabilidade, a topologia que apresenta melhor desempenho é a 2-20-12-2 uma vez que minimiza quer o DAM quer a MQE. A partir da análise da Tabela 27 pode afirmar-se que ambas as topologias apresentam valores de *bias* próximos de zero não apresentando, por isso, tendência para subestimar ou para sobrestimar as variáveis de saída.

Deste modo, para modelar os parâmetros nitrato e oxidabilidade nos sistemas de abastecimento de água para consumo humano do Distrito de Évora, a partir dos valores do pH e de condutividade, foram seleccionadas duas RNAs. No caso de se pretender maior acuidade na previsão do parâmetro nitrato dever-se-á utilizar a topologia 2-19-10-2, que consiste numa rede com uma camada de entrada com dois nós, correspondentes às variáveis condutividade e pH, duas camadas intermédias com, respectivamente, dezano-ve e dez nós e, por último, uma camada de saída com dois nós, correspondentes às variáveis nitrato e oxidabilidade (Figura 21 (a)). Se, pelo contrário, se pretender maior acuidade na previsão do parâmetro oxidabilidade dever-se-á utilizar a topologia 2-20-12-2

i.e. uma camada de entrada com dois nós (correspondentes às variáveis condutividade e pH), duas camadas intermédias com, respectivamente, vinte e doze nós e, por último, uma camada de saída com dois nós, correspondentes às variáveis nitrato e oxidabilidade (Figura 21 (b)).

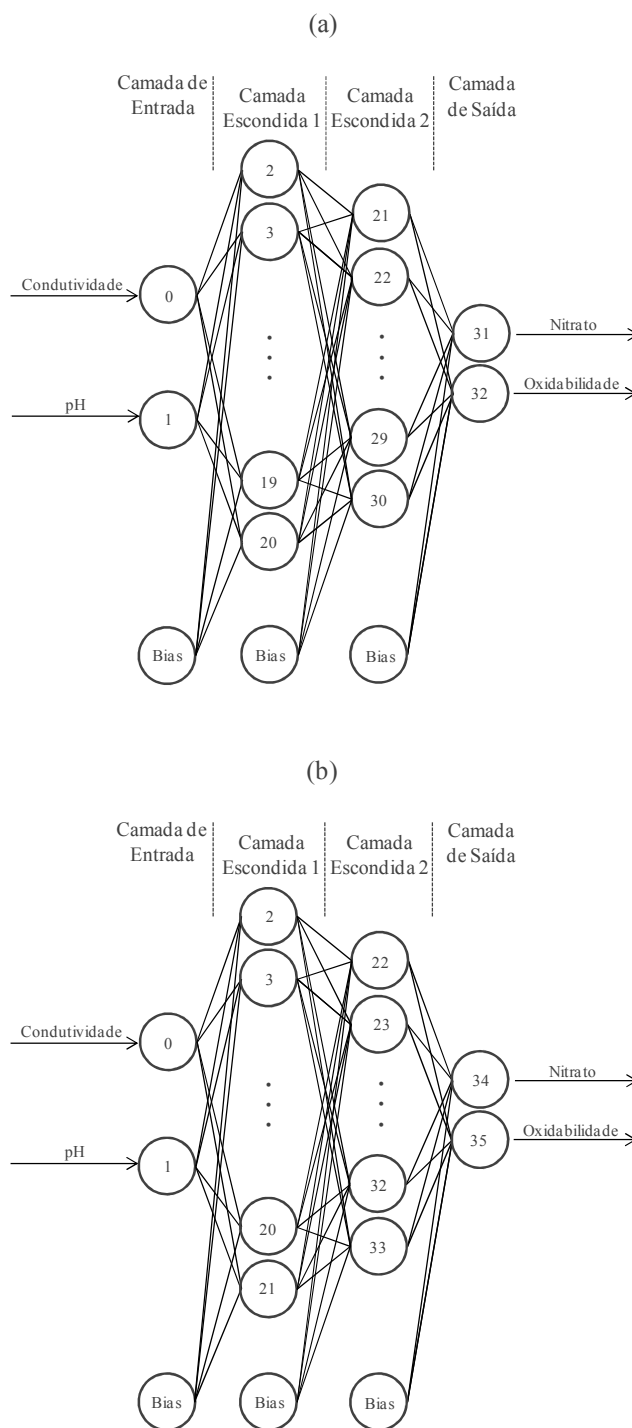


Figura 21 – RNAs seleccionadas para a previsão dos parâmetros nitrato e oxidabilidade a partir dos parâmetros condutividade e pH com as topologias 2-19-10-2 (a) e 2-20-12-2 (b)

Tendo em vista avaliar o desempenho dos modelos seleccionados apresentam-se as representações gráficas dos valores previstos em função dos valores experimentais para o conjunto de treino e para o conjunto de teste, tanto para a topologia 2-19-10-2 (Figura 22), como para a topologia 2-20-12-2 (Figura 23) indicando-se, para todas elas, os valores de R^2 . A análise dos gráficos apresentados na Figura 22 e na Figura 23, bem como dos respectivos valores de R^2 , mostram que a topologia 2-19-10-2 é a mais adequada para a previsão do parâmetro nitrato apresentando, porém, um pior desempenho na previsão do parâmetro oxidabilidade. Em contrapartida, a topologia 2-20-12-2 parece ser adequada para a previsão da oxidabilidade, a partir da condutividade e do pH, em todo o intervalo de valores ensaiado. Contudo, na previsão do parâmetro nitrato apresenta um desempenho inferior à da topologia 2-19-10-2, tal como se esperava a partir da análise dos valores do DAM e da MQE.

Os valores de R^2 , relativos aos parâmetros de saída, tanto para o conjunto de treino como para o conjunto de teste, são semelhantes apontando para a não existência de sobre-ajustamento em ambos os modelos.

Com o objectivo de verificar se os modelos não subestimam ou sobrestimam as variáveis dependentes, construíram-se os gráficos que se apresentam na Figura 24, onde se representa os resíduos em função dos valores previstos pelos modelos para ambas as variáveis de saída, tanto para o conjunto de treino como para o conjunto de teste.

A partir da análise quer da Figura 24 quer dos valores de R^2 que nela se apresentam, fica patente a distribuição aleatória dos resíduos, de ambos os lados da linha horizontal e o facto de serem independentes do valor da variável de saída. O intervalo de variação dos resíduos está de acordo com a ideia apresentada anteriormente que aponta para o facto de o modelo 2-19-10-2 apresentar um desempenho mais elevado na previsão dos valores de nitrato comparativamente com o modelo 2-20-12-2.

Tal como anteriormente foi referido, não é possível apresentar um conjunto independente de dados para validação do modelo. Contudo, é possível validar os modelos relativamente à sua capacidade para prever o parâmetro nitrato utilizando 134 casos (Tabela 13 e Tabela 24). Para o efeito, forneceram-se os valores dos parâmetros de entrada (condutividade e pH) e compararam-se os valores previstos pelos dois modelos seleccionados com os valores experimentais (Figura 25).

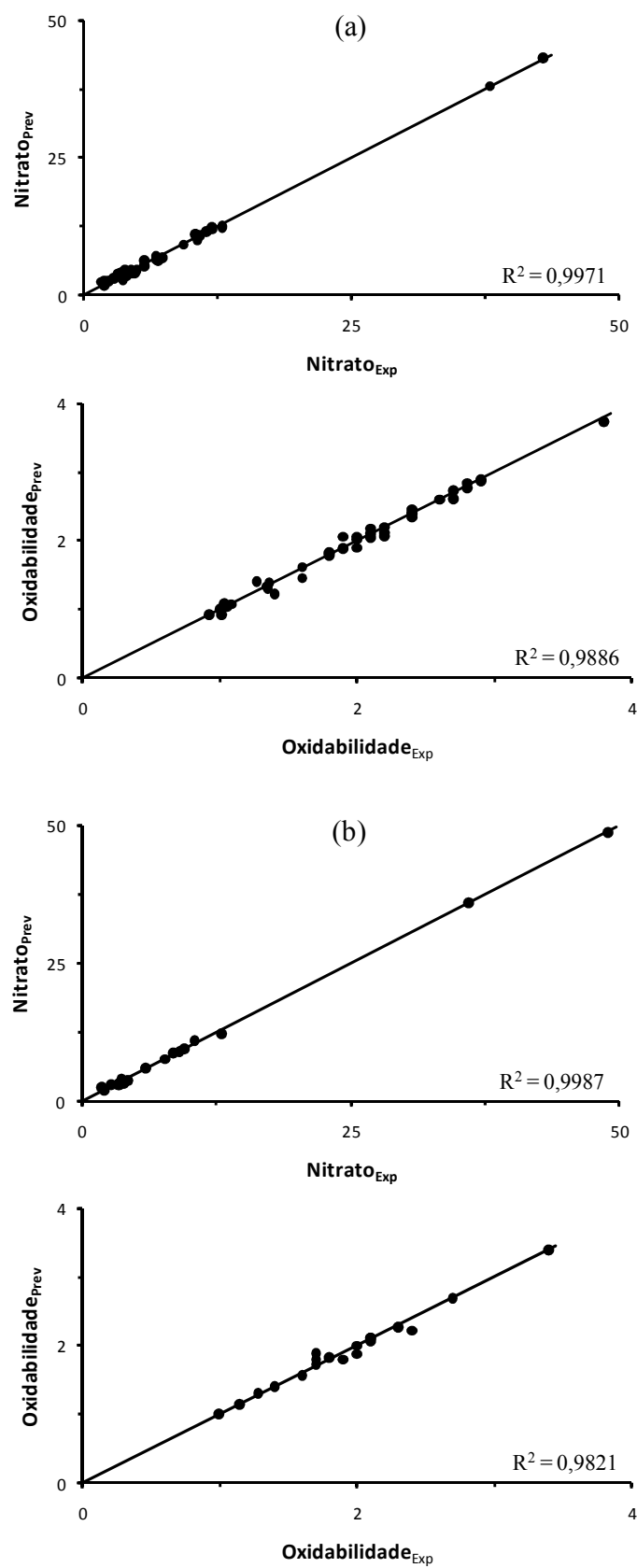


Figura 22 – Valores previstos vs valores experimentais das variáveis de saída, para o conjunto de treino (a) e para o conjunto de teste (b), com base na topologia 2-19-10-2

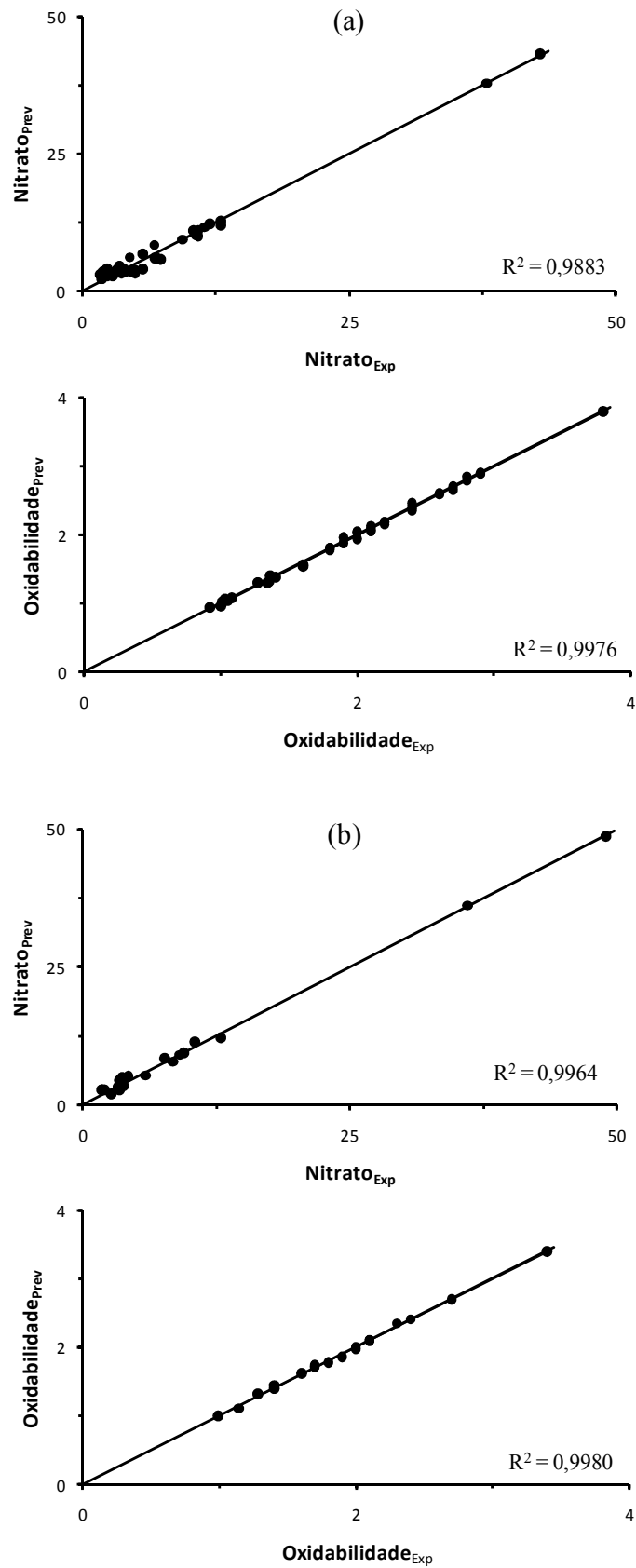


Figura 23 – Valores previstos vs valores experimentais das variáveis de saída, para o conjunto de treino (a) e para o conjunto de teste (b), com base na topologia 2-20-12-2

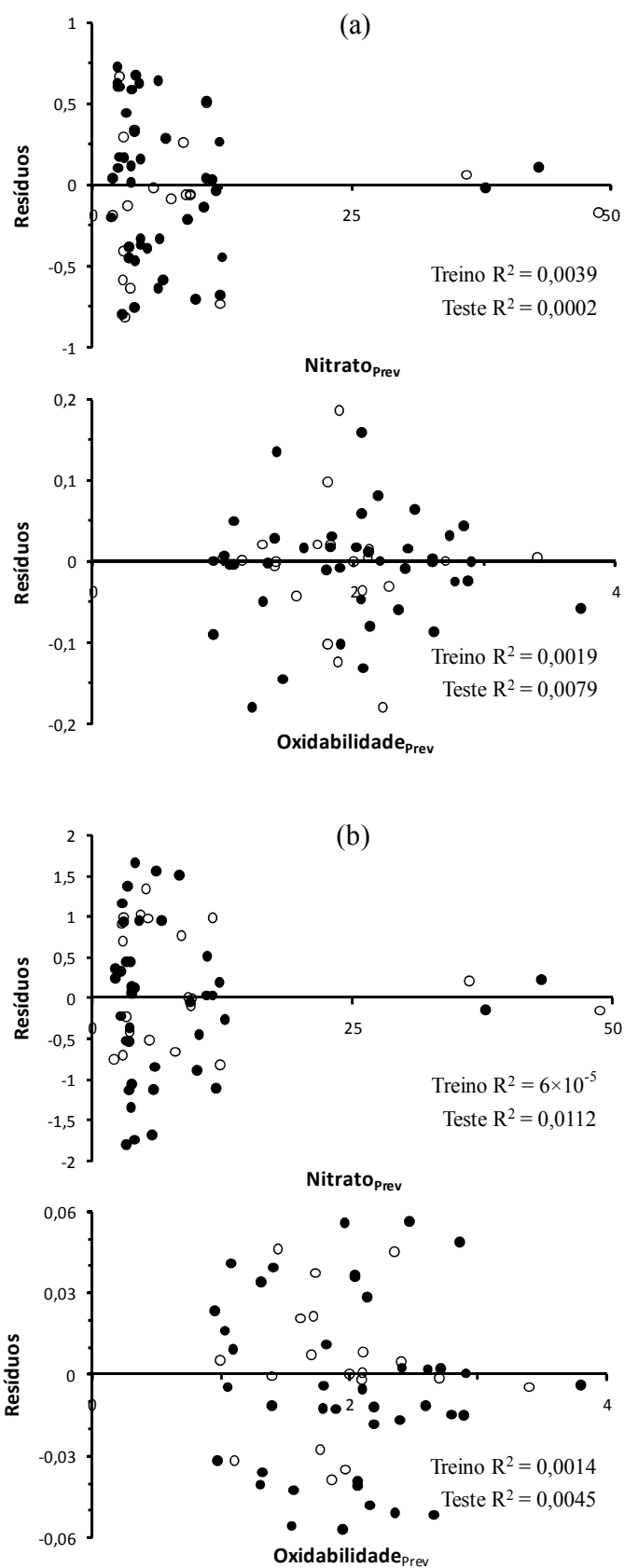


Figura 24 – Resíduos vs valores previstos para as variáveis de saída para os modelos com as topologias 2-19-10-2 (a) e 2-20-12-2 (b). (●) Conjunto de treino (○) conjunto de teste

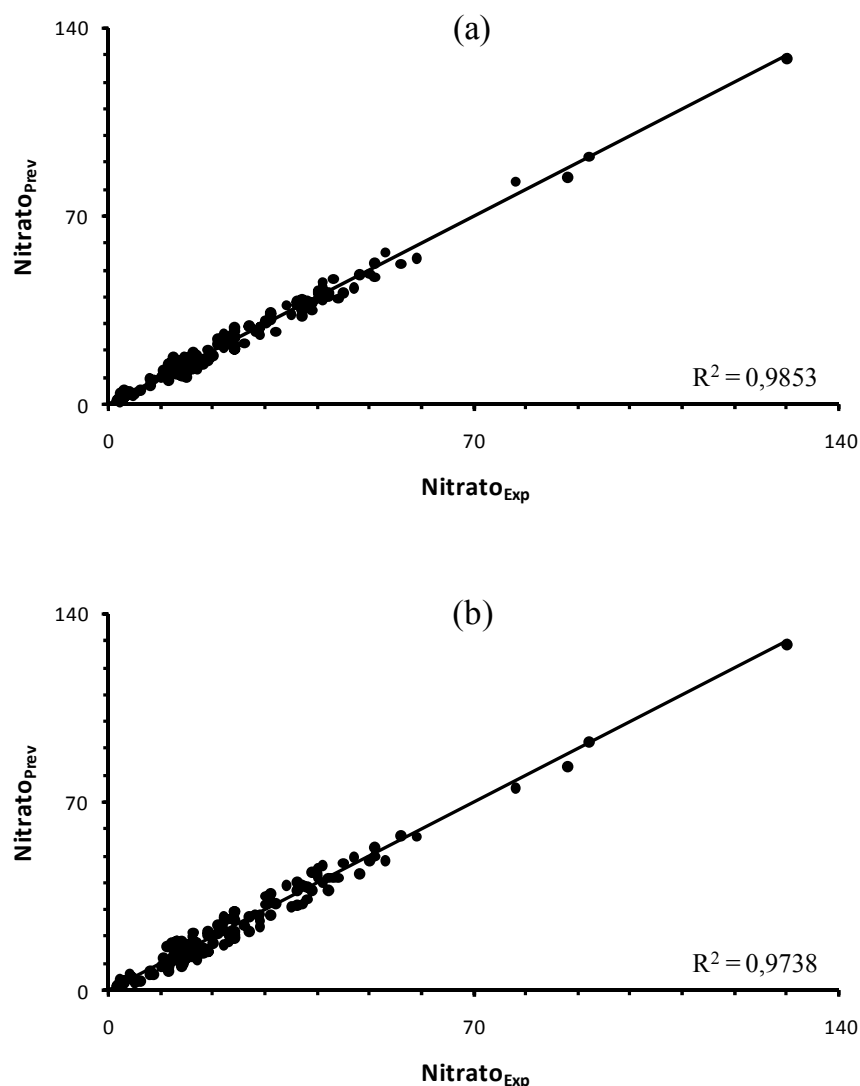


Figura 25 – Valores previstos vs valores experimentais relativos ao conjunto de validação para o parâmetro nitrato para modelos com as topologias 2-19-10-2 (a) e 2-20-12-2 (b)

A análise quer dos gráficos apresentados na Figura 25 quer dos valores de R^2 , que nela se apresentam, mostra que mesmo para dados independentes o desempenho do modelo 2-19-10-2, no que respeita à previsão do parâmetro nitrato, se mantém elevado sendo superior ao desempenho do outro modelo. Tal como anteriormente, para o conjunto de validação, apresentam-se as distribuições dos resíduos em função dos valores de nitrato previstos (Figura 26) verificando-se que são independentes dos valores das saídas e se distribuem aleatoriamente. Estes resultados parecem, pois, reforçar a afirmação anterior relativa ao facto de os modelos não apresentarem problemas de sobre-ajustamento.

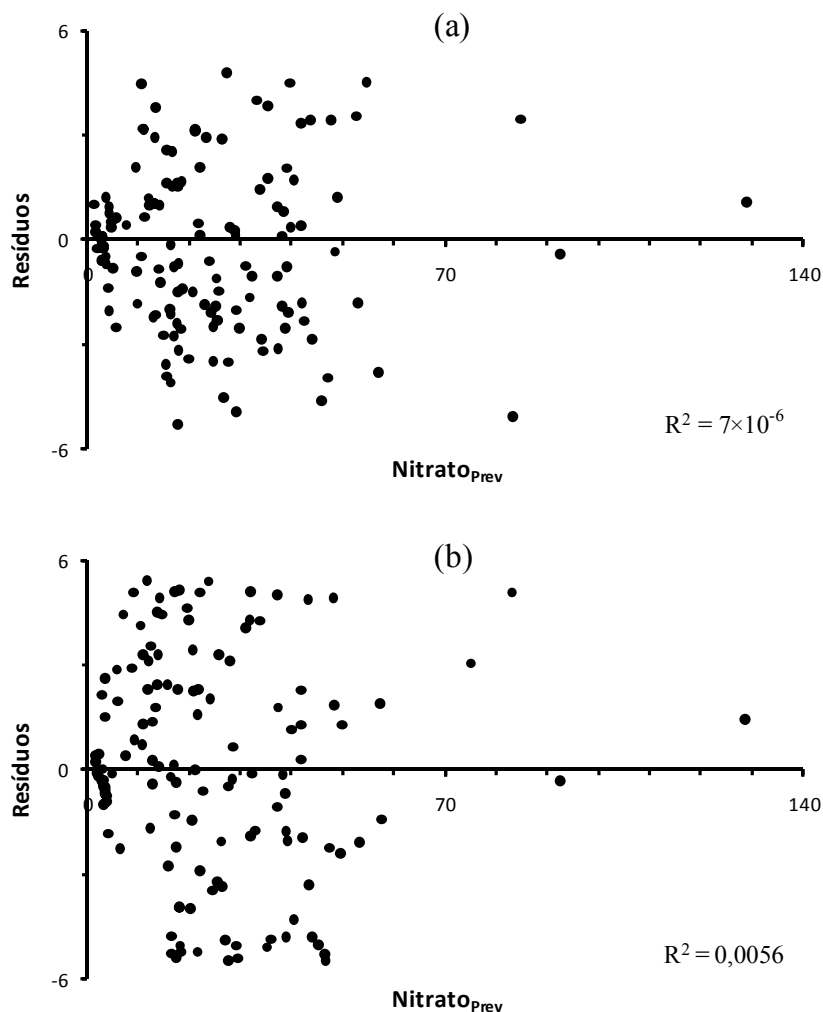


Figura 26 – Resíduos vs valores previstos relativos ao conjunto de validação para o parâmetro nitrato para os modelos com as topologias 2-19-10-2 (a) e 2-20-12-2 (b)

V.3.2.3 – Modelos de previsão dos parâmetros nitrato e manganês

Para a modelação dos parâmetros nitrato e manganês, com base nos valores de pH e da condutividade elaboraram-se e testaram-se algumas topologias de rede. Na Tabela 28 encontram-se os valores do DAM, da MQE e de *bias* para algumas das topologias testadas, tanto para o conjunto de treino como para o conjunto de teste. A sua análise mostra que a topologia 2-19-10-2 é aquela que minimiza tanto o DAM como a MQE revelando-se, de entre as topologias testadas, aquela que apresenta melhor desempenho na previsão dos teores em nitrato. No que se refere à previsão do parâmetro manganês, por seu lado, é a topologia 2-19-17-2 aquela que apresenta melhor desempenho uma vez que minimiza quer o DAM quer a MQE.

Tabela 28 – Valores do DAM, da MQE e de *bias* para algumas das topologias testadas

Topologia	Nitrato		Manganês	
	Treino	Teste	Treino	Teste
DESVIO ABSOLUTO MÉDIO (DAM)				
2-9-5-2	3,012	3,122	1,878	2,417
2-12-5-2	3,821	3,750	3,757	6,525
2-14-10-2	1,946	1,243	1,765	1,481
2-17-10-2	1,951	1,216	1,892	1,228
2-19-10-2	0,346	0,287	0,900	1,122
2-19-17-2	1,286	0,888	0,332	0,370
2-24-11-2	2,191	1,941	2,060	3,095
MÉDIA DO QUADRADO DOS ERROS (MQE)				
2-9-5-2	15,575	19,291	7,429	9,404
2-12-5-2	39,937	49,997	26,969	98,597
2-14-10-2	8,358	3,592	7,941	3,364
2-17-10-2	8,716	2,495	9,413	2,553
2-19-10-2	0,162	0,104	1,495	1,873
2-19-17-2	2,462	1,188	0,158	0,181
2-24-11-2	9,491	5,550	9,222	27,882
BIAS				
2-9-5-2	-1,092	-0,441	0,028	-0,207
2-12-5-2	0,692	-0,876	-0,058	0,103
2-14-10-2	-0,093	-0,453	-0,291	0,281
2-17-10-2	-0,494	0,522	-0,240	0,251
2-19-10-2	-0,018	0,043	-0,084	0,290
2-19-17-2	-0,153	-0,037	0,011	0,101
2-24-11-2	-0,531	0,170	-0,161	-0,279

A partir da análise da Tabela 28 pode afirmar-se, também, que ambas as topologias apresentam valores de *bias* próximos de zero, não apresentando, por isso, tendência para subestimar ou para sobrestimar as variáveis de saída.

Atendendo ao exposto no parágrafo anterior, foram seleccionadas duas RNAs para modelar o teor de nitrato e manganês nos sistemas de abastecimento de água para consumo humano do Distrito de Évora, a partir dos valores do pH e de condutividade. No caso de se pretender colocar a ênfase na previsão do parâmetro nitrato dever-se-á utilizar a topologia 2-19-10-2, que consiste numa rede com uma camada de entrada com dois nós, correspondentes às variáveis condutividade e pH, duas camadas intermédias com, respectivamente, dezanove e dez nós e, por último, uma camada de saída com dois nós, correspondentes às variáveis manganês e nitrato (Figura 27 (a)). Se, pelo contrário, o acento tónico for colocado na previsão do parâmetro manganês dever-se-á utilizar a topologia 2-19-17-2 i.e. uma camada de entrada com dois nós (correspondentes às variáveis condutividade e pH), duas camadas intermédias com, respectivamente, dezanove e dezassete nós e, por último, uma camada de saída com dois nós, correspondentes às variáveis manganês e nitrato (Figura 27 (b)).

No que concerne à avaliação do desempenho dos modelos seleccionados apresentam-se as representações gráficas dos valores previstos em função dos valores experimentais para o conjunto de treino e para o conjunto de teste, tanto para a topologia 2-19-10-2 (Figura 28), como para a topologia 2-19-17-2 (Figura 29) indicando-se, para todas elas, os valores do coeficiente de determinação.

A análise dos gráficos apresentados na Figura 28 e na Figura 29 bem com os valores de R^2 obtidos, mostram que o modelo com a topologia 2-19-10-2 parece ser adequado para a previsão do parâmetro nitrato, apresentando um pior desempenho na previsão do parâmetro manganês. O modelo com a topologia 2-19-17-2, pelo contrário, parece ser adequado para a previsão dos valores de manganês, a partir do conhecimento da condutividade e do pH, em todo o intervalo de valores ensaiado. Relativamente à previsão do parâmetro nitrato o desempenho é inferior ao do modelo 2-19-10-2, tal como se esperava a partir da análise dos valores do DAM e da MQE. O facto de os valores de R^2 relativos aos conjuntos de treino e de teste serem semelhantes, para cada um dos parâmetros de saída parece mostrar que ambos os modelos não apresentam problemas de sobre-ajustamento.

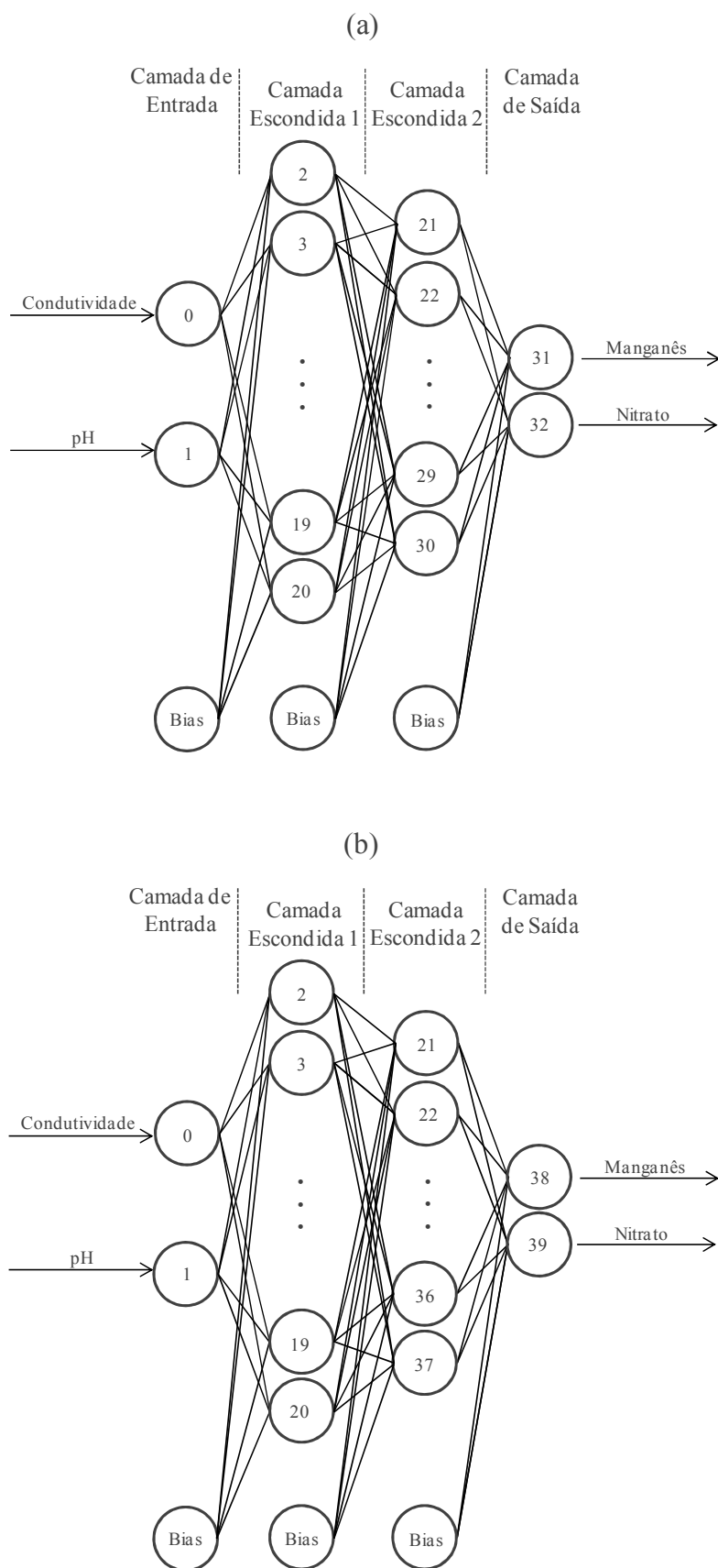


Figura 27 – RNAs seleccionadas para a previsão dos parâmetros manganês e nitrato a partir dos parâmetros condutividade e pH com as topologias 2-19-10-2 (a) e 2-19-17-2 (b)

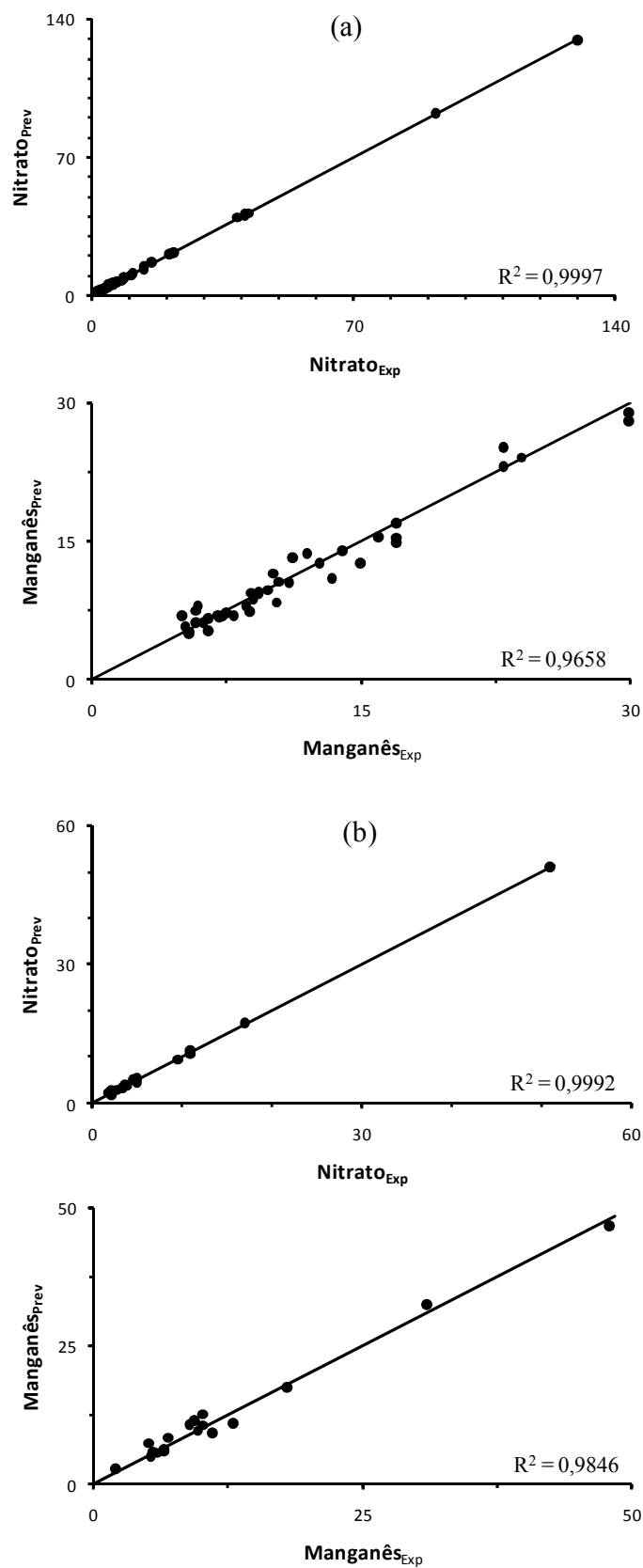


Figura 28 – Valores previstos vs valores experimentais das variáveis de saída, para o conjunto de treino (a) e para o conjunto de teste (b), com base na topologia 2-19-10-2

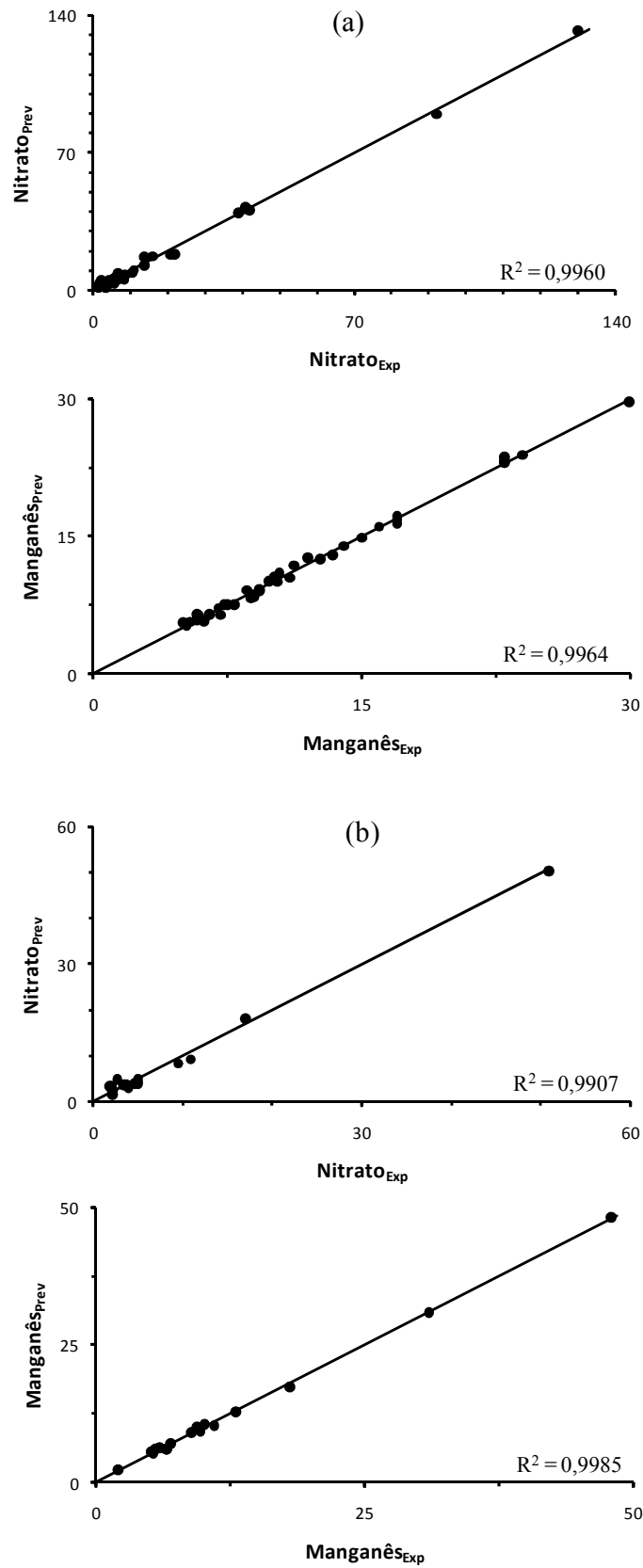


Figura 29 – Valores previstos vs valores experimentais das variáveis de saída, para o conjunto de treino (a) e para o conjunto de teste (b), com base na topologia 2-19-17-2

Como foi anteriormente salientado, os valores de *bias* parecem mostrar que os modelos seleccionados não apresentam tendência para subestimar ou sobrestimar as variáveis de saída. Para confirmar tal facto construíram-se os gráficos que se apresentam na Figura 30, onde se representou o resíduo em funções dos valores previstos pelos modelos para cada uma das variáveis de saída, tanto para o conjunto de treino como para o conjunto de teste.

A análise quer dos gráficos da Figura 30 quer dos valores de R^2 , que nela se apresentam, mostra que os resíduos se distribuem de modo aleatório, de ambos os lados da linha horizontal e são independentes do valor da variável de saída. O intervalo de variação dos resíduos acentua a ideia apresentada anteriormente, que o modelo 2-19-10-2 apresenta elevado desempenho na previsão dos valores de nitrato enquanto o modelo 2-19-17-2 é o mais adequado para prever os valores de manganês.

O número de registos existentes na BD com informação completa relativa aos parâmetros nitrato e manganês, tal como anteriormente, não permite ter um conjunto independente de dados para validar os modelos. Contudo, é possível validá-los relativamente à sua capacidade para prever os valores da concentração de nitrato utilizando 136 casos (Tabela 13 e Tabela 24). Para o efeito, forneceram-se os valores de pH e de condutividade e compararam-se os valores previstos pelos dois modelos seleccionados com os valores experimentais (Figura 31).

A análise dos gráficos da Figura 31 bem como dos valores de R^2 , que nela se apresentam, mostra que mesmo para dados independentes o desempenho do modelo 2-19-10-2, no que respeita à previsão do parâmetro nitrato, se mantém elevado sendo superior ao desempenho do outro modelo. A distribuição dos resíduos em função dos valores de nitrato previstos, para os dados de validação, para ambos os modelos, encontram-se na Figura 32 onde, tal como anteriormente, se verifica que são independentes dos valores das saídas e se distribuem aleatoriamente. Estes resultados parecem, pois, corroborar a afirmação anterior relativa ao facto de os modelos não apresentarem problemas de sobre-ajustamento.

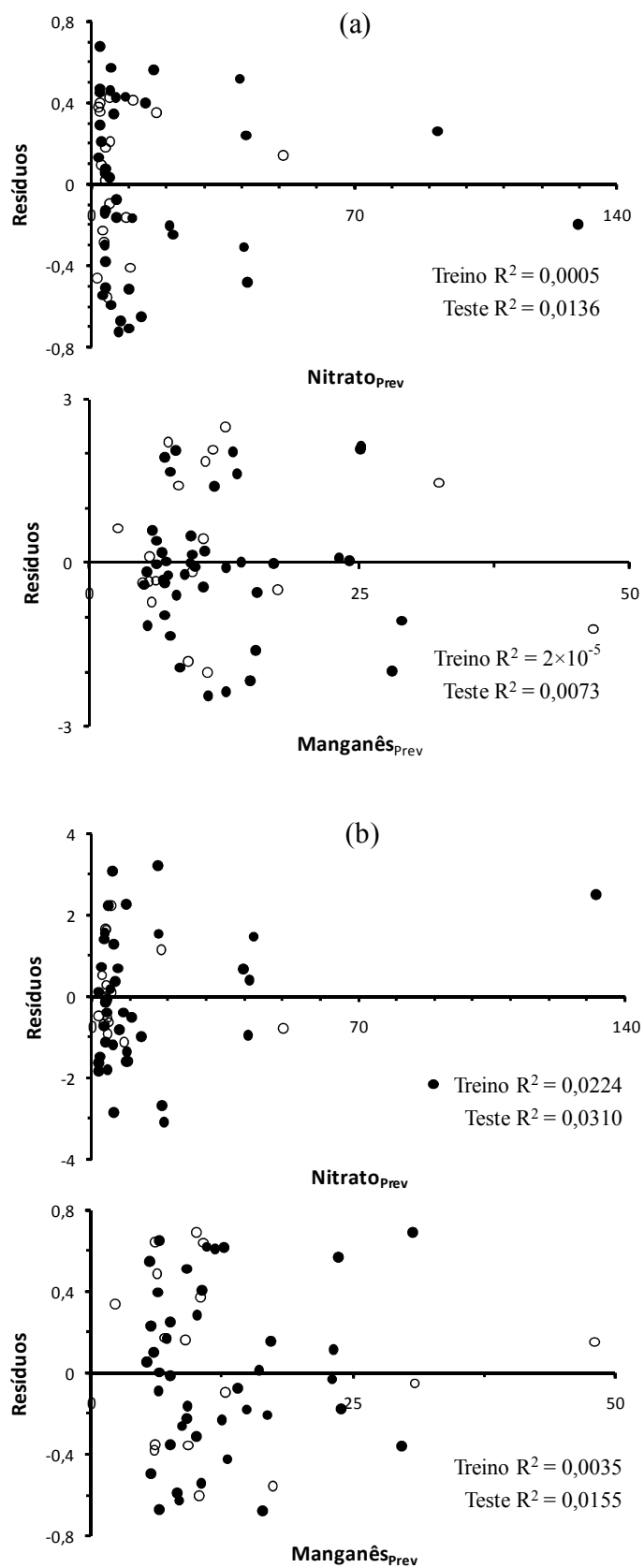


Figura 30 – Resíduos vs valores previstos para as variáveis de saída para os modelos com as topologias 2-19-10-2 (a) e 2-19-17-2 (b). (●) Conjunto de treino (○) conjunto de teste

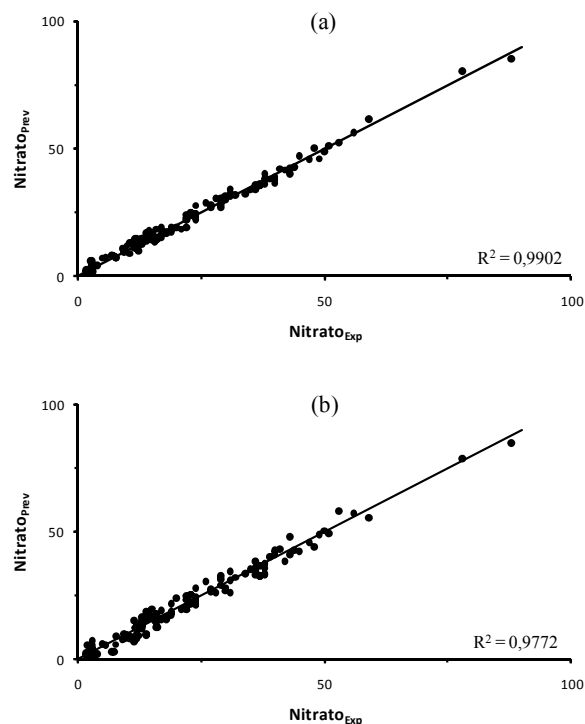


Figura 31 – Valores previstos vs valores experimentais relativos ao conjunto de validação para o parâmetro nitrato para modelos com as topologias 2-19-10-2 (a) e 2-19-17-2 (b)

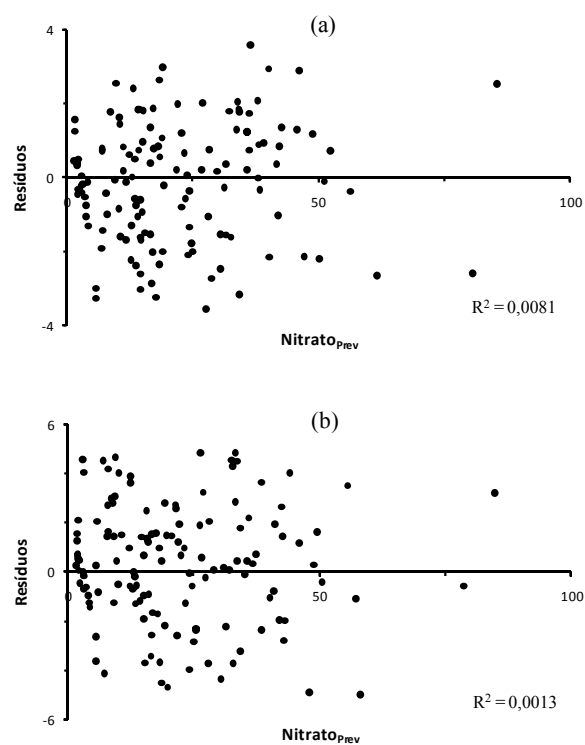


Figura 32 – Resíduos vs valores previstos relativos ao conjunto de validação para o parâmetro nitrato para os modelos com as topologias 2-19-10-2 (a) e 2-19-17-2 (b)

V.3.2.4 – Modelos de previsão dos parâmetros nitrato e turvação

Para a modelação dos parâmetros nitrato e turvação, com base nos valores de pH e da condutividade elaboraram-se e testaram-se algumas topologias de rede. Na Tabela 29 encontram-se os valores do DAM, da MQE e de *bias* para algumas das topologias testadas, tanto para o conjunto de treino como para o conjunto de teste.

Tabela 29 – Valores do DAM, da MQE e de *bias* para algumas das topologias testadas

Topologia	Nitrato		Turvação	
	Treino	Teste	Treino	Teste
DESVIO ABSOLUTO MÉDIO (DAM)				
2-5-5-2	2,748	2,577	0,230	0,230
2-7-4-2	2,028	1,768	0,038	0,035
2-7-6-2	2,280	2,206	0,234	0,147
2-8-6-2	0,964	0,844	0,123	0,173
2-10-6-2	2,810	2,368	0,123	0,206
2-13-8-2	2,021	2,251	0,213	0,253
2-14-8-2	1,929	1,834	0,170	0,300
MÉDIA DO QUADRADO DOS ERROS (MQE)				
2-5-5-2	12,145	10,253	0,124	0,117
2-7-4-2	6,175	5,018	0,002	0,002
2-7-6-2	7,369	9,957	0,127	0,033
2-8-6-2	1,433	1,037	0,038	0,098
2-10-6-2	15,440	10,943	0,038	0,084
2-13-8-2	6,765	10,155	0,120	0,174
2-14-8-2	5,673	6,879	0,076	0,187

Tabela 29 (Cont.) – Valores do DAM, da MQE e de bias para algumas das topologias testadas

Topologia	Nitrato		Turvação	
	Treino	Teste	Treino	Teste
<i>BIAS</i>				
2-5-5-2	-0,443	0,184	-0,078	-0,049
2-7-4-2	-0,438	-0,200	-0,008	-0,010
2-7-6-2	-0,421	-0,202	-0,066	-0,030
2-8-6-2	-0,323	-0,472	-0,072	-0,097
2-10-6-2	-0,844	0,683	-0,035	-0,009
2-13-8-2	0,050	-0,312	-0,052	-0,019
2-14-8-2	0,110	0,144	-0,055	0,003

A análise da Tabela 29 mostra que a topologia 2-8-6-2 é aquela que minimiza tanto o DAM como a MQE revelando-se, de entre as topologias testadas, aquela que apresenta melhor desempenho na previsão dos teores em nitrato. No que se refere à previsão da turvação a topologia 2-7-4-2 é aquela que apresenta melhor desempenho, uma vez que minimiza quer o DAM quer a MQE. A partir da análise da Tabela 29 pode afirmar-se que ambas as topologias apresentam valores de *bias* próximos de zero, não apresentando, assim, tendência para subestimar ou para sobrestimar as variáveis de saída.

Atendendo ao exposto no parágrafo anterior, foram seleccionadas duas RNAs para modelar o teor de nitrato e turvação nos sistemas de abastecimento de água para consumo humano do Distrito de Évora, a partir dos valores do pH e de condutividade. No caso de se pretender colocar o acento tónico na previsão do parâmetro nitrato dever-se-á utilizar a topologia 2-8-6-2, que consiste numa rede com uma camada de entrada com dois nós, correspondentes às variáveis condutividade e pH, duas camadas intermédias com, respectivamente, oito e seis nós e, por último, uma camada de saída com dois nós, correspondentes às variáveis nitrato e turvação (Figura 33 (a)). Se, pelo contrário, a ênfase for colocada na previsão do parâmetro turvação dever-se-á utilizar a topologia 2-7-4-2 i.e. uma camada de entrada com dois nós (correspondentes às variáveis condutividade e pH), duas camadas intermédias com, respectivamente, sete e

No que se refere à avaliação do desempenho dos modelos seleccionados apresentam-se as representações gráficas dos valores previstos em função dos valores experimentais para o conjunto de treino e para o conjunto de teste, tanto para a topologia 2-8-6-2 (Figura 34), como para a topologia 2-7-4-2 (Figura 35) indicando-se, para todas elas, os valores do coeficiente de determinação. A análise dos referidos gráficos bem como os valores de R^2 obtidos, mostram que o modelo com a topologia 2-8-6-2 parece ser adequado para a previsão do parâmetro nitrato, apresentando um pior desempenho na previsão do parâmetro turvação. O modelo com a topologia 2-7-4-2, pelo contrário, parece ser mais adequado para a previsão dos valores de turvação, a partir do conhecimento da condutividade e do pH, em todo o intervalo de valores ensaiado. Relativamente à previsão do parâmetro nitrato o desempenho é inferior ao do modelo 2-8-6-2, tal como se esperava a partir da análise dos valores do DAM e da MQE.

O facto de os valores de R^2 relativos aos conjuntos de treino e de teste serem semelhantes, para cada um dos parâmetros de saída parece mostrar que ambos os modelos não apresentam problemas de sobre-ajustamento.

Como foi anteriormente destacado, os valores de *bias* parecem mostrar que os modelos seleccionados não apresentam tendência para subestimar ou sobrestimar as variáveis de saída. Para o confirmar construíram-se os gráficos que se apresentam na Figura 36, onde se representou o resíduo em função dos valores previstos pelos modelos para cada uma das variáveis de saída, tanto para o conjunto de treino como para o conjunto de teste.

A análise quer dos gráficos da Figura 36 quer dos valores de R^2 , que nela se apresentam, mostra que os resíduos se distribuem de modo aleatório, de ambos os lados da linha horizontal e são independentes do valor da variável de saída. O intervalo de variação dos resíduos reforça a ideia, já apresentada, que o modelo 2-8-6-2 apresenta elevado desempenho na previsão dos valores de nitrato enquanto o modelo 2-7-4-2 é o mais adequado para prever os valores de turvação.

Tendo em vista validar os modelos relativamente à sua capacidade para prever os valores da concentração de nitrato para os 144 casos não utilizados (Tabela 13 e Tabela 24), forneceram-se os valores de pH e de condutividade e compararam-se os valores previstos pelos modelos seleccionados com os valores experimentais (Figura 37).

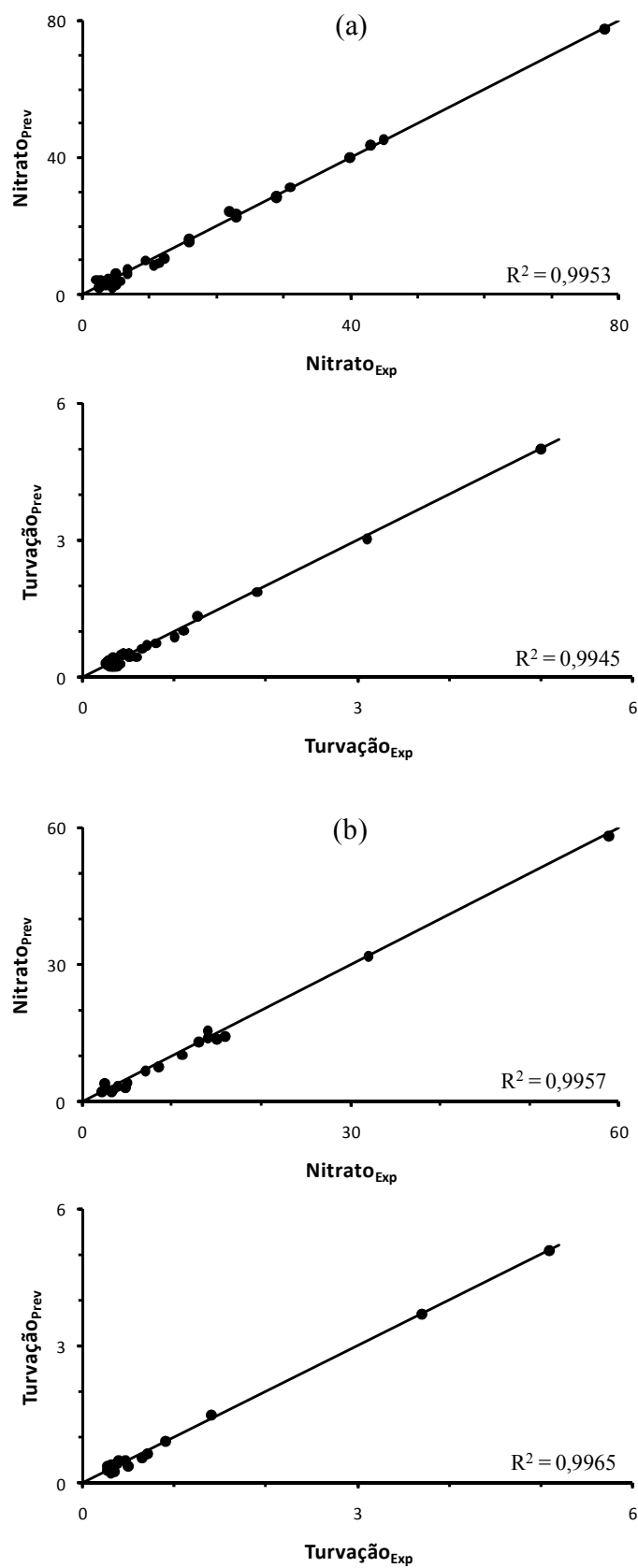


Figura 34 – Valores previstos vs valores experimentais das variáveis de saída, para o conjunto de treino (a) e para o conjunto de teste (b), com base na topologia 2-8-6-2

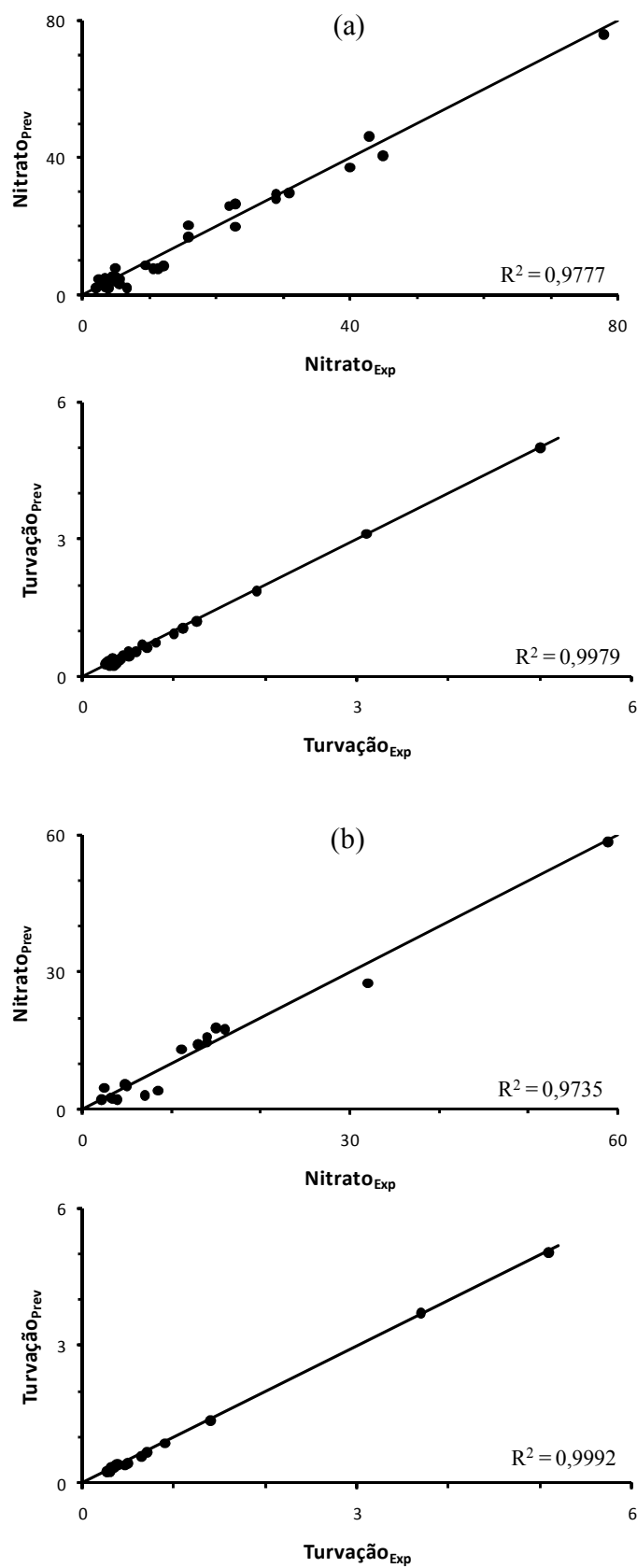


Figura 35 – Valores previstos vs valores experimentais das variáveis de saída, para o conjunto de treino (a) e para o conjunto de teste (b), com base na topologia 2-7-4-2

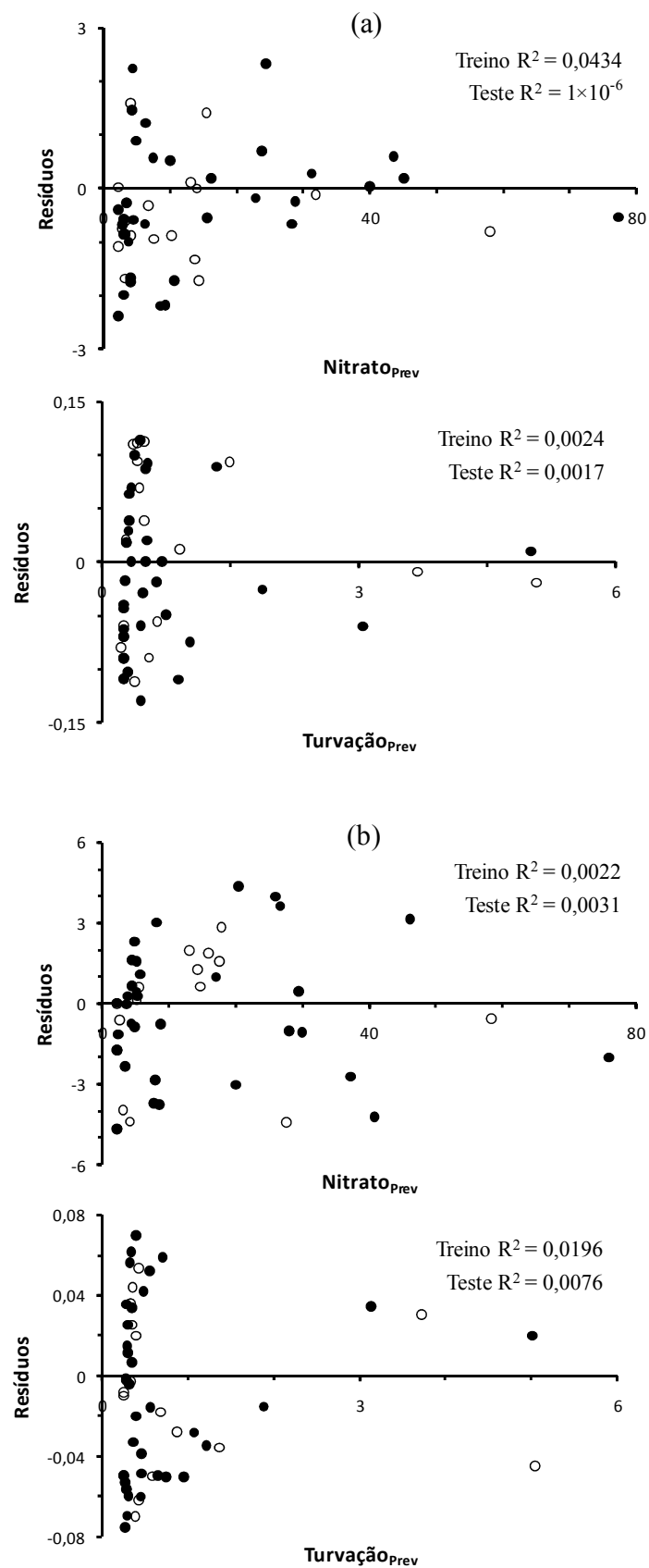


Figura 36 – Resíduos vs valores previstos para as variáveis de saída para os modelos com as topologias 2-8-6-2 (a) e 2-7-4-2 (b). (●) Conjunto de treino (○) conjunto de teste

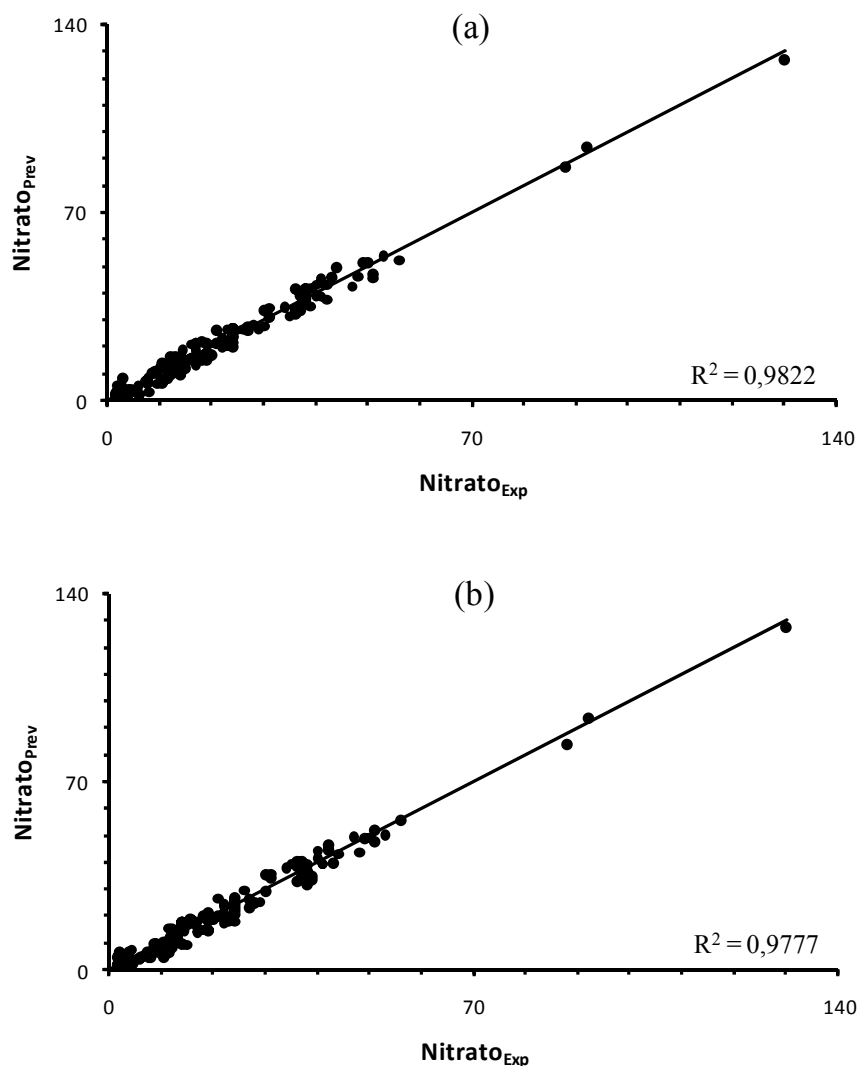


Figura 37 – Valores previstos vs valores experimentais relativos ao conjunto de validação para o parâmetro nitrato para modelos com as topologias 2-8-6-2 (a) e 2-7-4-2 (b)

A análise dos gráficos da Figura 37 bem como dos valores de R^2 , que nela se apresentam, mostra que o desempenho do modelo 2-8-6-2, no que respeita à previsão do parâmetro nitrato, se mantém elevado sendo superior ao desempenho do outro modelo. A distribuição dos resíduos em função dos valores de nitrato previstos, para os dados de validação, para ambos os modelos, encontram-se na Figura 38 onde, tal como anteriormente, se verifica que são independentes do valor da variável de saída e se distribuem aleatoriamente. Estes resultados parecem, pois, indicar que os modelos não apresentam problemas de sobre-ajustamento.

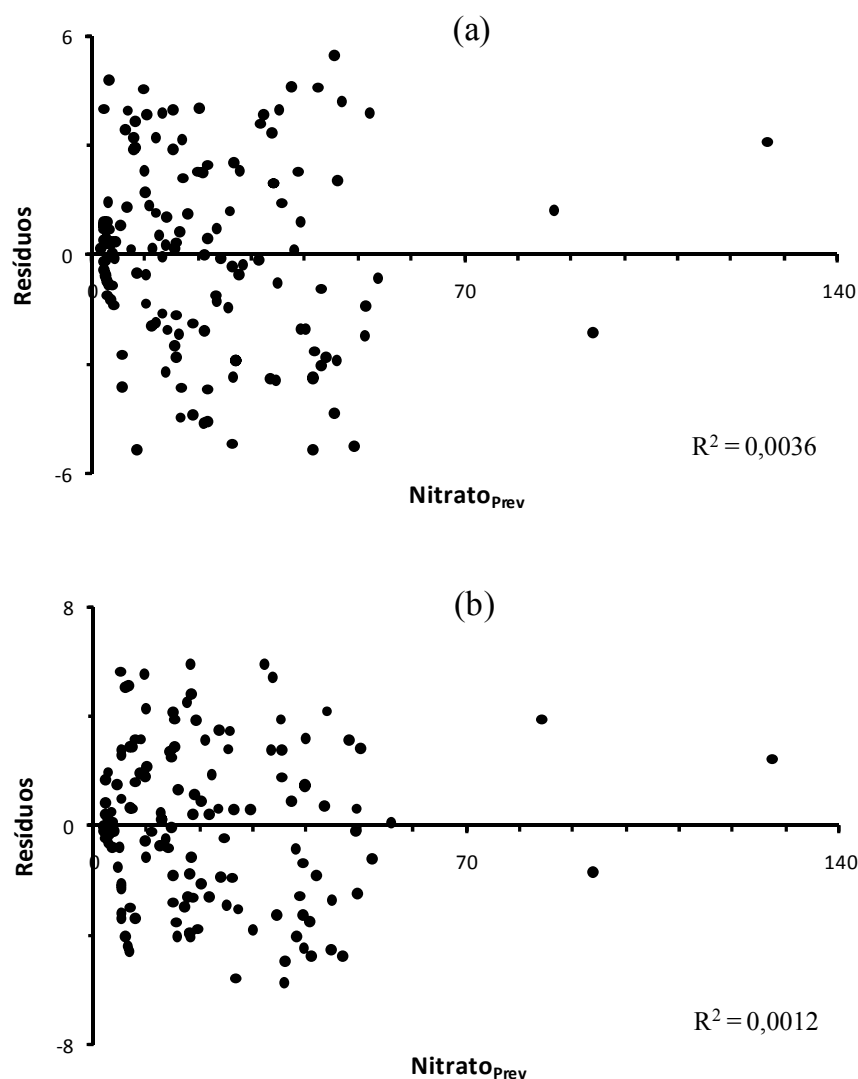


Figura 38 – Resíduos vs valores previstos relativos ao conjunto de validação para o parâmetro nitrato para os modelos com as topologias 2-8-6-2 (a) e 2-7-4-2 (b)

V.3.2.5 – Análise de Sensibilidade dos Modelos Apresentados

De um modo geral, pretende-se concentrar os esforços na aquisição dos dados referentes às variáveis que mais contribuem para a acuidade do modelo. A análise de sensibilidade é um procedimento que permite analisar o efeito das variações nas entradas sobre as variáveis de saída, através do cálculo da importância relativa de cada uma das variáveis de entrada.

A importância das variáveis de entrada pode ser feita, por exemplo, com base nos pesos das conexões entre as várias camadas que constituem a rede [Lee et al., 2003], ou com

base numa análise de gradiente [Kewley et al., 2000]. Neste trabalho, contudo, optou-se pela análise de sensibilidade mais comumente utilizada que consiste na determinação da variância produzida na saída (V_x) quando todas as variáveis de entrada são mantidas no seu valor médio excepto uma, que varia ao longo da sua escala de valores [Kewley et al., 2000]:

$$V_x = \frac{\sum_{i=1}^N (\hat{Y}_i - \bar{\hat{Y}})^2}{N - 1} \quad (15)$$

onde \hat{Y}_i denota o valor da saída obtido com todas as variáveis de entrada mantidas no seu valor médio excepto a variável x e $\bar{\hat{Y}}$ denota o valor da saída obtido com todas as variáveis de entrada mantidas no seu valor médio.

A importância relativa (IR_x) das variáveis utilizadas pelo modelo é dada pelo quociente entre a variância produzida na saída pela variável x e o somatório das variâncias produzidas na saída por todas as variáveis de entrada:

$$IR_x = \frac{V_x}{\sum_j V_{xj}} \quad (16)$$

Na Figura 39 apresenta-se a importância relativa das variáveis condutividade e pH para todos os modelos apresentados anteriormente.

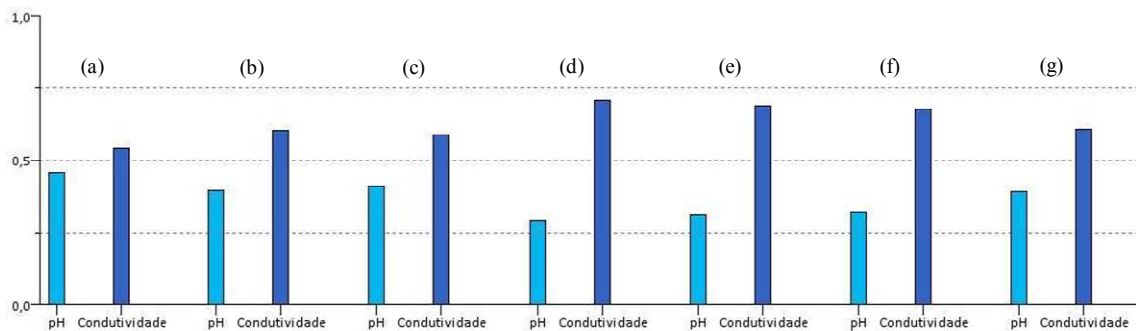


Figura 39 – Importância relativa das variáveis de entrada para os modelos desenvolvidos para previsão do nitrato, potássio e sódio com a topologia 2-18-14-3 (a); nitrato e oxidabilidade com as topologias 2-19-10-2 (b) e 2-20-12-2 (c); nitrato e manganês com as topologias 2-19-10-2 (d) e 2-19-17-2 (e); nitrato e turvação com as topologias 2-8-6-2 (f) e 2-7-4-2 (g)

A análise do gráfico apresentado na Figura 39 mostra que em todos os modelos a importância relativa da variável condutividade é superior à da variável pH, variando entre 0,54 e 0,71. Tal facto parece indicar que o referido parâmetro é aquele que mais influencia as variáveis de saída em particular nos modelos referentes à previsão de nitrato e manganês (d) e (e).

O resultado apresentado no parágrafo anterior pode estar relacionado com o facto de a condutividade expressar, de um modo geral, a concentração de iões presentes na água que, para muitas espécies, é dependente do pH. Deste modo, a informação sobre o parâmetro pH pode estar, de forma indirecta, contida na condutividade. O facto da importância relativa do pH ser maior nos modelos que envolvem sódio e potássio (a) e oxidabilidade (b) e (c) parece corroborar a afirmação anterior, uma vez que se trata ou de espécies muito solúveis ou de uma medida da matéria orgânica onde a influência do pH não se deve repercutir nos valores da condutividade.

VI – CONCLUSÕES FINAIS E PERSPECTIVAS FUTURAS

Após a conclusão do trabalho, é chegado o momento de olhar para trás, analisar o percurso feito e perspectivar o trabalho futuro.

O presente trabalho dá uma contribuição para a resolução de problemas relacionados com a qualidade da água, tendo-se recorrido a técnicas e metodologias da Área Científica da Inteligência Artificial e da Descoberta de Conhecimento em Bases de Dados utilizando, como ferramenta, as Redes Neurais Artificiais. Desenvolveram-se modelos de previsão de alguns parâmetros dos grupos de análises P_2 e P_3 com base nos parâmetros físico-químicos relativos ao grupo de análises P_1 , que se caracterizam por apresentarem desempenhos elevados, exibindo coeficientes de determinação superiores a 0,965, 0,980 e 0,970, respectivamente para os conjuntos de treino, teste e validação.

Atendendo a que, na maior parte dos casos, são pedidas ao laboratório análise do grupo P_1 sendo as análises dos grupos P_2 e P_3 feitas de uma forma mais espaçada, os modelos desenvolvidos permitem aferir a qualidade global da água e potenciar a previsão, dando resposta ao principal desafio lançado constituindo-se, assim, como exemplo da aplicação das ferramentas e técnicas de *Data Mining* à resolução de problemas na área da saúde pública.

Tendo em vista compreender a influência do local de amostragem procurou-se extrair conhecimento dos dados físico-químicos, através de métodos que utilizassem a aprendizagem não-supervisionada, tendo recorrido a modelos de segmentação utilizando a estratégia *k-means*. Os modelos obtidos revelam diferenças entre as propriedades físico-químicas das águas dos diversos concelhos analisados permitindo, após a construção de modelos explicativos da segmentação, caracterizar e separar as amostras provenientes de origem subterrânea das amostras relativas a águas de origem superficial.

O trabalho elaborado e apresentado na presente dissertação abre espaço para o desenvolvimento de ferramentas automáticas de suporte à decisão com recurso à programação em lógica estendida que se prevê que venha a reforçar respostas tendo em

vista, por um lado, melhorar a qualidade da água disponibilizada às populações e, por outro, contribuir para um aproveitamento sustentado dos recursos hídricos.

O trabalho futuro que se perspectiva visa a criação de um Universo Virtual e Dinâmico, constituído por entidades complexas e sociabilizáveis tais como genes, genomas, organismos e populações. Tais entidades complexas e interactivas competem entre si, estando sujeitas a regimes de selecção, tendo em vista modelar e predizer a qualidade da água disponibilizada à população.

Os cenários de evolução do sistema, serão obtidos recorrendo à Computação Genética e Evolucionária, cujo regime de selecção assentará sobre um processo de quantificação da qualidade da informação emanada de um programa em lógica que descreve o sistema, em que o ambiente computacional será construído em torno de um processo de Raciocínio Baseado em Casos.

No final, uma agência para a integração, difusão e arquivo de dados irá emergir, a qual será utilizada para alicerçar um protótipo de um Sistema Multi-Agente para Modelação e Simulação da Qualidade da Água de Abastecimento Público, que se possa afirmar como um Sistema de Apoio à Decisão e possa contribuir para a tomada da melhor decisão em tempo útil.

REFERÊNCIAS BIBLIOGRÁFICAS

- Adriaans, P. & Zantinge, D., Data Mining. Addison-Wesley, Edinburgh, U.K., 1996.
- Alley, E. R., Water Quality Control Handbook. 2ª Ed., McGraw-Hill, New York, U.S.A., 2007.
- Almeida, C., Mendonça, J., Jesus, M. & Gomes, A., Sistemas Aquíferos de Portugal Continental. INAG, Lisboa, Portugal, 2000.
- Apte, C., Liu, B., Pednault, E. & Smith, P., Business applications of Data Mining. *Communications of the ACM*, 45: 49-53: 2002.
- Arbib, A. Eds., The Handbook of Brain Theory and Neural Networks. 2ª Ed, MIT Press, Cambridge, U.K., 2003.
- ARS-Norte, Significado dos Parâmetros Incluídos na Vertente Analítica do Programa de Vigilância Sanitária da Água para Consumo Humano. Administração Regional de Saúde do Norte - I.P., Ministério da Saúde, Lisboa, Portugal. 2008.
- ARSA lentejo, <http://www.arsalentejo.min-saude.pt>, Acedido em 8 de Julho, 2010.
- Atkins, J. P., Burdon, D. & Allen, J. H., An application of contingent valuation and decision tree analysis to water quality improvements. *Marine Pollution Bulletin*, 55: 591-602, 2007.
- Berry, M. & Linoff, G., Mastering Data Mining: The Art and Science of Customer Relationships Management. John Wiley & Sons, New York, U.S.A., 2000.
- Birrien, J. I., História da Informática. Rés Editora, Porto, Portugal, 2008.
- Bose, N. & Liang, P., Neural Networks Fundamentals with Graphs, Algorithms and Applications. McGraw-Hill Education, Maidenhead, U.K., 1996.
- Boser, B., Guyon, I. & Vapnik, V., A training algorithm for optimal margin classifiers. *Proceedings of 5th Annual Workshop on Computational Learning Theory*, 144-152, ACM, Burlington, U.S.A., 1992.
- Bradley, P. S. & Fayyad, U. M., Refining Initial Points for K-Means Clustering. Em J. Shavlik Eds., *Proceedings of 15th International Conference on Machine Learning (ICML98)*, 91- 99, Morgan Kaufmann, San Francisco, U.S.A., 1998.
- Bramer, M., Principles of Data Mining. Springer-Verlag, London, U.K., 2007.

- Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J., Classification and Regression Trees. Chapman & Hall / CRC Press, Boca Raton, U.S.A., 1984.
- Broomhead, D. & Lowe, D., Multivariable functional interpolation and adaptive networks. *Complex Systems*, 2: 321-355, 1988.
- Butenko, S., Chaovalitwongse, W. A. & Pardalos, P. M., Clustering Challenges in Biological Networks. World Scientific, New Jersey, U.S.A., 2009.
- Chakrabarti, S., Cox, E., Frank, E., Güting, R. H., Han, J., Jiang, X., Kamber, M., Lighstone, S. S., Nadeau, T., Neapolitan, R., Pyle, D., Refaat, M., Schneider, M., Teorey, T. & Witten, I., Data Mining: Know It All. Morgan Kaufmann, Burlington, U.S.A., 2009.
- Chang, G., Healey, M., McHugh, J. & Wang, J., Mining the Word Wide Web - An Information Search Approach. Kluwer Academic Publishers, Massachusetts, U.S.A., 2001.
- Chenard, J.-F. & Caissie, D., Stream temperature modelling using artificial neural networks: application on Catamaran Brook. *Hydrological Processes*, 22: 3361–3372, 2008.
- Cheng, B. & Titterton, D. M., Neural networks: a review from a statistical perspective. *Statistical Science*, 9: 2-30, 1994.
- Cios, K. J., Pedrycz, W., Swiniarski, R. W. & Kurgan, L. A., Data Mining Knowledge Discovery Approach. Springer. New York, U.S.A, 2007.
- Clapham, D., Small Water Supplies. 2ª Ed., Clay's Library of Health and the Environment, New York, U.S.A., 2004.
- Clementine, Clementine SPSS – User Guide version 11.1, Integral Solution limited. 2007.
- Cortez, P., Aprendizagem e Avaliação de Modelos – Apontamentos Pedagógicos. Departamento de Sistemas de Informação, Universidade do Minho, Guimarães, Portugal, 2004.
- Costa, E. & Simões, A., Inteligência Artificial - Fundamentos e Aplicações. 1ª Ed., FCA - Editora de Informática Lda., Lisboa, Portugal, 2008.
- Couto, C., Avaliação e Controlo da Água de Albufeiras com Base em Modelos Dinâmicos, Dissertação de Mestrado, Universidade de Évora, Évora, Portugal, 2010.
- Cruz, A., Data Mining via Redes Neurais Artificiais e Máquinas de Vectores de Suporte, Dissertação de Mestrado, Universidade de Minho, Braga, Portugal, 2007.
- Dogan, E., Sengorur, B. & Koklu, R., Modeling biochemical oxygen demand of the Melen River in Turkey using an artificial neural network technique. *Journal of Environmental Management*, 90: 1229–1235, 2009.

- Eaton, A., Clesceri, L., Rice, E. & Greenberg, A. Eds., Standard Methods for the Examination of Water and Wastewater. 21^a Ed., American Public Health Association, Washington, U.S.A., 2005.
- Fayyad, U., Mining Databases: Towards Algorithms for Knowledge Discovery. *IEEE Computer Society Technical Committee on Data Engineering*, 21: 39-48, 1997.
- Fayyad, U., Piatetshy-Shapiro, G. & Uthurusamy, R., Advances in Knowledge Discovery and Data Mining. AAAI Press/MIT Press, Massachussetts, U.S.A, 1996a.
- Fayyad, U., Piatetshy-Shapiro, G., Uthurusamy, R. & Smyth, P., From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, 17: 37-54, 1996b.
- Fernandes, A. T., Previsão da Classificação da Qualidade dos Efluentes da Indústria Petrolífera com Recurso a Sistemas Inteligentes, Dissertação de Mestrado, Universidade de Évora, Évora, Portugal, 2009.
- Fernandes, A. V., Vicente, H. & Neves, J., Solving Challenging Problems in the Oil Industry Using Artificial Intelligence Based Tools. Em Diganta B. Das, V. Nassehi & L. Deka Eds., *Proceedings of Industrial Simulation Conference 2009 - ISC' 2009*, 325-331, Eurosis - ETI Publication, Loughborough, U.K., 2009.
- Flood, I. & Kartam, N., Neural network in civil engineering: I. Principles and understanding. *Journal of Computational in Civil Engineering*, 8: 131-148, 1994.
- Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B., Bayesian Data Analysis. 2^a Ed., Chapman & Hall / CRC, Boca Raton, U.S.A., 2003.
- Goebel, M. & Gruenwald, L., A survey of Data Mining and knowledge discovery software tools. *ACM SIGKDD Explorations*, 1: 20-33, 1999.
- Gopalan, N. P. & Sivaselvan, B., Data Mining: Techniques and Trends. PHI Learning, New Delhi, India, 2009.
- Govindaraju, R. S. & Rao, A. R., Artificial Neural Networks in Hydrology. Kluwer Academic Publishers, Dordrecht, Holanda, 2010.
- Gray, N. F., Drinking Water Quality - Problems and Solutions. Cambridge University Press, Cambridge, U.K., 2008.
- Groth, R., Data Mining: Building Competitive Advantage. Prentice Hall PTR, New Jersey, U.S.A., 2000.
- Hagan, M. T., Demuth, H. B. & Beale, M., Neural Network Design. PWS Publishing Company, Boston, U.S.A., 1996.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. & Witten, I. H., The WEKA Data Mining Software: An Update. *SIGKDD Exploration*, 11: 10-18, 2009.

- Han, J. & Kamber, M., Data Mining: Concepts and Techniques. 2^a Ed., Morgan Kauffmann Publishers, San Francisco, U.S.A., 2006.
- Hand, D., Mannila, H. & Smith, P., Principles of Data Mining. MIT Press, Cambridge, U.K., 2001.
- Hastie, T., Tibshirani, R. & Friedman, J., The Elements of Statistical Learning: Data Mining, Inference and Prediction. Springer-Verlag, New York, U.S.A., 2001.
- Haykin, S., Neural Networks and Learning Machines. 3^a Ed., Prentice Hall, New Jersey, U.S.A., 2008.
- Hebb, D., The Organization of Behavior: A Neuropsychological Theory. John Wiley & Sons Inc., New York, U.S.A., 1949.
- Hopfield, J., Neural networks and physical systems with emergent collective computational abilities. *The National Academy of Science*, 79: 2554-2558, 1982.
- Hornik, K., Stinchcombe, M. & White, H., Multilayer feed-forward networks are universal approximators. *Neural Networks*, 2: 359-366, 1989.
- Huheey, J. E., Keiter, E. A. & Keiter, R. L., Inorganic Chemistry: Principles of Structure and Reactivity. 4^a Ed., Prentice Hall, New Jersey, U.S.A., 1997.
- Jordan, M., Why the logistic function? A tutorial discussion on probabilities and neural networks. Computational Cognitive Science, Technical Report 9503, MIT, 1995.
- Kartalopoulos, S. V., Understanding Neural Network and Fuzzy Logic - Basic Concepts and Applications. IEEE Press, New Jersey, U.S.A, 1996.
- Karul, C., Soyupak, S., Çilesiz, A., Akbay, N. & Germen, E., Case studies on the use of neural networks in eutrophication modelling. *Ecological Modelling*, 134: 145-152, 2000.
- Kewley, R., Embrechts, M. & Breneman, C., Data strip mining for the virtual design of pharmaceuticals with neural networks. *IEEE Transactions on Neural Networks*, 11: 668-679, 2000.
- Kohavi, R. & Provost, F., Glossary of Terms. *Machine Learning*, 30: 271-274, 1998.
- Kohonen, T., Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43: 59-69, 1982.
- Kufrin, R., Generating C4.5 Productions Rules in Parallel; 4th National Conference on Artificial Intelligence (AAAI-97). AAAI Press, Rhode Island, U.S.A., 1997.
- Kuo, J.-T., Hsieh, M.-H., Lung, W.-S. & She, N., Using artificial neural network for reservoir eutrophication prediction. *Ecological Modelling*, 200: 171-177, 2007.
- Larose, D. T., Data Mining Methods and Models. John Wiley & Sons, Hoboken, U.S.A., 2006.

- Lee, J. H., Huang, Y., Dickman, M. & Jayawardena, A. W., Neural network modelling of coastal algal blooms. *Ecological Modelling*, 159: 179–210, 2003.
- Lehr, J. H., Keeley, J. & Lehr, J. Eds., Water Encyclopedia: Water Quality and Resource Development. 1ª Ed., John Wiley & Sons, Hoboken, U.S.A., 2005.
- Little, R., Regression with missing X's: a review. *Journal of the American Statistical Association*, 87: 1227-1237, 1992.
- Maimon, O. & Rokach, L., Decomposition Methodology for Knowledge Discovery and Data Mining: Theory and Applications. World Scientific Publishing, London, U.K., 2005.
- Mamedov, A. & Udalov, S., A computer tool to develop individual based models for simulation of population interactions. *Ecological Modelling*, 147: 53-68, 2002.
- Matoso, A., Rasga, M., Santana, M. & Murteira, M., Principais albufeiras do Alentejo monitorizadas. Comissão de Coordenação da Região Alentejo, Direcção de Serviços de Monitorização Ambiental, Évora, Portugal. 2004.
- McCulloch, W. & Pitts, W., A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5: 115-133, 1943.
- Mendes, B. & Oliveira, F. J., Qualidade da Água para Consumo Humano. Lidel - Edições Técnicas, Lisboa, Portugal, 2004.
- Michalski, R., Bratko, I. & Miroslav, K., Machine Learning and Data Mining Methods and Applications. John Willey & Sons, London, U.K., 1998.
- Midões, C. & Costa, A., Sistema Aquífero Estremoz-Cano. Em: Os Aquíferos das Bacias Hidrográficas do Rio Tejo e das Ribeiras do Oeste - Saberes e Reflexões, Administração da Região Hidrográfica do Tejo I.P., Ministério do Ambiente e Ordenamento do Território, Lisboa, Portugal. 2010.
- Minsky, M. & Papert, S., Perceptrons: An Introduction to Computational Geometry. MIT Press, Massachussets, U.S.A., 1969.
- Morgan, J. N. & Sonquist, J. A., Problems in the analysis of survey data, and a proposal. *Journal of the American Statistical Association*, 58: 415-434, 1963.
- Nisbet, R., Elder, J. & Miner, G., Handbook of Statistical Analysis & Data Mining Applications. Academic Press, San Diego, U.S.A., 2009.
- NIST/SEMATECH, e-Handbook of Statistical Methods.
<http://www.itl.nist.gov/div898/handbook>, Acedido em 20 de Setembro, 2010.
- OMS, <http://www.who.int/en/>, Acedido em 13 de Abril, 2010.
- Palani, S., Liong, S. & Tkalic, P., An ANN application for water quality forecasting. *Marine Pollution Bulletin*, 56: 1586-1597, 2008.

- Patterson, D., Artificial Neural Networks - Theory and Application. Prentice Hall, New York, U.S.A., 1996.
- Perlovsky, L. I., Neural Networks and Intellect: Using Model-Based Concepts. Oxford University Press, Oxford, U.K., 2001.
- Pinto, A., Fernandes, A. V., Vicente, H. & Neves, J., Optimizing Water Treatment Systems Using Artificial Intelligence Based Tools. Em C. A. Brebbia e V. Popov Eds., Water Resource Management V - WIT Transactions on Ecology and the Environmental, Vol. 125, pp. 185-194, WIT Press, Southampton, U.K., 2009.
- Pinto, A. M., Aplicação de Sistemas Inteligentes à Optimização de um Sistema de Tratamento de Águas, Dissertação de Mestrado, Universidade de Évora, Évora, Portugal, 2009.
- Popek, E., Sampling and analysis of environment chemical pollutants. Academic Press, London, U.K., 2006.
- Pyle, D., Data Preparation for Data Mining. Morgan Kaufmann, San Francisco, U.S.A., 1999.
- Quinlan, J. R., Simplifying decision trees. *International Journal of Man-Machine Studies*, 27: 221-23, 1987.
- Quinlan, J. R., C4.5: Programs for machine learning. Morgan Kaufmann Publishers Inc., San Mateo, U.S.A., 1993.
- Quinlan, J. R., Bagging, boosting and C4.5. Em Eds. *Proceedings of AAAI'96 National Conference on Artificial Intelligence*, 725-730, 1996.
- Rich, E. & Knight, K., Artificial Intelligence. 2ª Ed., McGraw-Hill Inc., New York, U.S.A., 1991.
- Rojas, R., Neural Networks - A Systematic Introduction. Springer-Verlag, Berlin, Alemanha, 1996.
- Rosenblatt, F., The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review*, 65: 386-408, 1958.
- Rubin, D. B., Multiple Imputation after 18 years. *Journal of the American Statistical Association*, 91: 473-489, 1996.
- Rumelhart, D., Hinton, G. & Williams, R., Learning Internal Representation by Error Propagation. Em David E. Rumelhart & James L. McClelland Eds., Parallel Distributed Processing, Vol. 1: Foundations, pp. 318-362, MIT Press, Massachusetts, U.S.A., 1986.
- Russell, S. J. & Norvig, P., Artificial Intelligence: A Modern Approach 2ª Ed., Prentice Hall, New Jersey, U.S.A., 2003.

- Santos, M. & Azevedo, C., Data Mining - Descoberta de Conhecimento em Base de Dados. FCA - Editora de Informática Lda, Lisboa, Portugal, 2005.
- Santos, M. F., Cortez, P., Quintela, H., Neves, J., Vicente, H. & Arteiro, J., Ecological Mining - A Case Study on Dam Water Quality. Em A. Zanasi, C.A. Brebbia & N.F.F. Ebecken Eds., Data Mining VI - Data Mining, Text Mining and their Business Applications, WIT Transactions of Information and Communication Technologies, Vol. 35, pp. 523-531, WIT Press, Southampton, U.K., 2005.
- Schafer, J., Analysis of Incomplete Multivariate Data. Chapman & Hall / CRC, Boca Raton, U.S.A., 1999.
- Schapire, E. R., The boosting approach to machine learning: An overview. *Proceedings of MSRI Workshop on Nonlinear Estimation and Classification*, 1-23, 2002.
- Singh, K. P., Basant, A., Malik, A. & Jain, G., Artificial neural network modeling of the river water quality - A case study. *Ecological Modelling*, 220: 888-895, 2009.
- SNIRH, <http://www.snirh.pt>, Acedido em 15 de Setembro, 2010.
- Soman, K. P., Diwakar, S. & Ajay, V., Data Mining: Theory and Practice. PHI Learning, New Delhi, India, 2006.
- Sousa, C., Data Mining: Metodologias, Tecnologias, Modelos e Aplicações. Dissertação de Mestrado, Universidade do Minho. Guimarães, Portugal, 2004.
- Soyupak, S., Karaer, F. & Gurbuz, H., A neural network based approach for calculating dissolved oxygen profiles in reservoirs. *Neural Computing & Applications*, 12: 166-172, 2003.
- Spellman, F. R., The Science of Water - Concepts and Applications. 2ª Ed., CRC Press, Boca Raton, U.S.A., 2008.
- Stumm, W. & Morgan, J. J., Aquatic Chemistry - Chemical Equilibria and Rates in Natural Waters. 3ª Ed., John Wiley & Sons, Inc., New York, U.S.A., 1996.
- Tanir, D. E., Data Mining and Knowledge Discovery Technologies. IGI Publishing, New York, U.S.A., 2008.
- Teixeira, H. Q., Sistemas de Conhecimento Baseados em Data Mining: Aplicação à Análise da Estabilidade de Estruturas Metálicas, Dissertação de Mestrado, Universidade do Minho, Braga, Portugal, 2005.
- Thuraisingham, B., Data Mining Technologies, Techniques, Tools and Trends. CRC Press LLC, Boca Raton, U.S.A., 1999.
- Torgo, L., Árvores de Regressão – Métodos e Aplicações, Tese de Doutoramento, LIACC, Faculdade de Engenharia do Porto, Porto, Portugal, 1999.
- Turban, E., Aronson, J. E. & Liang, T.-P., Decision Support Systems and Intelligent Systems. 7ª Edição, Prentice Hall, New Jersey, U.S.A., 2004.

- Varis, O., Tortajada, C. & Biswas, A. K. Eds., Water Resources Development and Management: Management of Transboundary Rivers and Lakes. Springer, Berlin, Alemanha, 2008.
- Vicente, H., Especificação e Prototipação de Sistemas de Gestão e Controlo da Qualidade da Água de Albufeiras, Departamento de Química, Universidade de Évora, Évora, Portugal, 2004.
- Weiner, E. R., Applications of Environmental Aquatic Chemistry - A Practical Guide. 2^a Ed., CRC Press, Boca Raton, U.S.A., 2008.
- Weiss, S. & Kulikowski, C., Computer Systems that Learn: Classification and Prediction Methods from Statistics, Neural Nets, Machine Learning, and Expert Systems. Morgan Kaufmann Publishers Inc., San Mateo, U.S.A., 1990.
- Williams, G. J. & Simoff, S. J., Data Mining: Theory, Methodology, Techniques, and Applications Springer-Verlag, Berlin, Alemanha, 2006.
- Witten, I. H. & Frank, E., Data Mining - Practical Machine Learning Tools and Techniques. 2^a Ed, Elsevier, Burlington, U.S.A., 2005.
- Zhang, C., Fundamentals of Environmental Sampling and Analysis. John Wiley & Sons, Inc., Hoboken, U.S.A., 2007.

ANEXO

Neste anexo apresenta-se a base de dados utilizada neste trabalho, referente às análises de água efectuadas no Distrito de Évora. <Lq denota inferior ao limite de quantificação do método utilizado e nd denota que a análise não foi pedida, não sendo o parâmetro determinado.

Concelho	Localidade/Freguesia	mês	Ano	pH	Condutividade	Amónio	Nitrato	Nitrito	Ferro	Turvação	Oxidabilidade	Potássio	Sódio	Manganês	Cádmio	Crómio	Cobre	Níquel	Chumbo	Alumínio
Alandroal	Alandroal	2	2007	7,4	590	<Lq	20	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Alandroal	Alandroal	9	2008	7,4	570	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Alandroal	2	2009	7,3	550	<Lq	12,3	<Lq	<Lq	0,32	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Alandroal	Alandroal	5	2009	7,4	610	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Alandroal	7	2007	7,2	620	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Alandroal	9	2007	7,3	510	<Lq	11,6	<Lq	<Lq	<Lq	<Lq	0,73	240	<Lq	<Lq	<Lq	0,013	<Lq	<Lq	nd
Alandroal	Alandroal	1	2008	7,6	580	<Lq	14	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Alandroal	Alandroal	5	2008	7,4	550	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Alandroal	10	2009	7,3	530	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Alandroal	11	2009	7,8	530	<Lq	11,4	<Lq	<Lq	0,33	<Lq	1,2	43	<Lq	<Lq	<Lq	0,05	<Lq	<Lq	nd
Alandroal	Alandroal	11	2007	7,4	570	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Aldeia de Plas	9	2008	8	670	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Aldeia de Plas	10	2009	8	570	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Aldeia de Plas	11	2009	8,1	530	<Lq	11,1	<Lq	<Lq	0,9	<Lq	1,2	49	6,2	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Alandroal	Aldeia de Plas	2	2007	8	530	<Lq	14	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Alandroal	Aldeia de Plas	5	2007	8	600	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Aldeia de Plas	9	2007	7,9	550	<Lq	8,6	<Lq	<Lq	<Lq	<Lq	nd	nd	16	nd	nd	nd	nd	nd	nd
Alandroal	Aldeia de Plas	5	2009	7,4	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Aldeia de Plas	11	2007	8,1	550	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Aldeia de Plas	5	2008	8	560	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Aldeia de Plas	12	2008	8	520	<Lq	12	<Lq	<Lq	<Lq	<Lq	1,1	25	<Lq	<Lq	<Lq	0,02	<Lq	<Lq	nd
Alandroal	Aldeia de Plas	7	2007	7,8	610	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Juromenha	7	2007	7,3	650	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Juromenha	2	2007	7,3	610	<Lq	44	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Alandroal	Juromenha	5	2007	7	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Juromenha	9	2007	7,2	660	<Lq	38	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Alandroal	Juromenha	11	2007	7,2	660	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Juromenha	5	2008	7,2	670	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Juromenha	9	2008	7,3	650	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Juromenha	5	2009	7,2	650	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Juromenha	10	2009	7,4	680	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Terena	2	2007	7,8	540	<Lq	14	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Alandroal	Terena	9	2007	7,6	480	<Lq	11,6	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Alandroal	Terena	5	2007	7,7	550	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Terena	11	2007	7,8	530	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Terena	5	2008	7,8	530	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Terena	9	2008	7,8	540	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Terena	5	2009	7,5	530	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Alandroal	Terena	10	2009	7,5	520	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd

Concelho	Localidade/Freguesia	mês	Ano	pH	Condutividade	Amónio	Nitrato	Nitrito	Ferro	Turvação	Oxidabilidade	Potássio	Sódio	Manganês	Cádmio	Crómio	Cobre	Níquel	Chumbo	Alumínio
Arraiolos	Aldeia da Serra	4	2007	8	760	<Lq	41	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Arraiolos	Aldeia da Serra	4	2009	7,6	730	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Aldeia da Serra	7	2009	7,7	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Aldeia da Serra	2	2007	8	760	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Aldeia da Serra	7	2007	8	780	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Aldeia da Serra	11	2007	8,1	760	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Aldeia da Serra	4	2008	8,1	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Aldeia da Serra	7	2008	8,1	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Arraiolos	1	2007	7,7	600	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Arraiolos	3	2007	7,6	610	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Arraiolos	3	2007	7,1	580	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Arraiolos	3	2007	7,6	580	<Lq	16	<Lq	<Lq	0,58	<Lq	nd	nd	7	nd	nd	nd	nd	nd	nd
Arraiolos	Arraiolos	4	2007	7,8	590	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Arraiolos	6	2007	7,7	600	<Lq	17	<Lq	<Lq	<Lq	<Lq	nd	nd	5,5	nd	nd	nd	nd	nd	nd
Arraiolos	Arraiolos	7	2007	7,6	550	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Arraiolos	9	2007	7,8	570	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Arraiolos	10	2007	7,7	560	<Lq	17	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Arraiolos	Arraiolos	3	2008	7,8	590	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Arraiolos	5	2008	7,7	600	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Arraiolos	11	2008	7,8	610	<Lq	16,2	<Lq	<Lq	<Lq	<Lq	2,5	46	<Lq	1,6	<Lq	<Lq	<Lq	8,2	nd
Arraiolos	Arraiolos	2	2009	7,8	590	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Arraiolos	6	2009	7,8	600	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Arraiolos	11	2009	7,9	900	<Lq	39	<Lq	<Lq	<Lq	<Lq	4,3	100	5,4	<Lq	<Lq	0,03	<Lq	<Lq	nd
Arraiolos	Igrejinha	9	2007	7,5	710	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Igrejinha	3	2008	7,7	730	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Igrejinha	3	2007	7,4	780	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Igrejinha	5	2007	7,4	770	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Igrejinha	11	2007	7,6	770	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Igrejinha	2	2007	7,5	780	<Lq	53	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Arraiolos	Igrejinha	4	2007	7,6	780	<Lq	50	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Arraiolos	Igrejinha	7	2007	7,4	810	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Igrejinha	10	2008	7,7	690	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Igrejinha	2	2009	7,3	770	<Lq	51	<Lq	<Lq	<Lq	<Lq	3,1	38	<Lq	<Lq	<Lq	0,007	<Lq	<Lq	nd
Arraiolos	Igrejinha	10	2009	7,6	810	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	São Gregório	2	2007	7,6	780	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	São Gregório	4	2007	7,7	820	<Lq	29	<Lq	<Lq	3,1	<Lq	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	São Gregório	6	2007	7,6	790	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	São Gregório	9	2007	7,6	720	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	São Gregório	11	2008	7,6	730	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	São Gregório	2	2009	7,2	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	São Gregório	11	2009	7,5	720	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	S. Pedro da Gafanhoeira	6	2007	7,3	377	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	S. Pedro da Gafanhoeira	8	2007	7,2	417	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	S. Pedro da Gafanhoeira	2	2007	7	385	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	S. Pedro da Gafanhoeira	3	2007	7	379	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	S. Pedro da Gafanhoeira	3	2007	7,4	384	<Lq	3,2	<Lq	90	0,32	1,03	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Arraiolos	S. Pedro da Gafanhoeira	7	2007	7,6	560	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	S. Pedro da Gafanhoeira	3	2008	7	383	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	S. Pedro da Gafanhoeira	4	2008	7	385	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	S. Pedro da Gafanhoeira	7	2008	6,9	422	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	S. Pedro da Gafanhoeira	4	2009	6,9	418	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	S. Pedro da Gafanhoeira	7	2009	7,1	478	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	S. Pedro da Gafanhoeira	5	2007	7,1	375	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Sabugueiro	3	2007	7,6	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Sabugueiro	4	2007	7,9	413	<Lq	nd	<Lq	200	16,6	3,7	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Sabugueiro	6	2007	7,9	810	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Sabugueiro	8	2007	8	770	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd

Concelho	Localidade/Freguesia	mês	Ano	pH	Condutividade	Amónio	Nitrato	Nitrato	Ferro	Turvação	Oxidabilidade	Potássio	Sódio	Manganês	Cádmio	Crómio	Cobre	Níquel	Chumbo	Alumínio
Arraiolos	Sabugueiro	6	2008	7,3	384	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Sabugueiro	9	2009	7,6	580	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Sabugueiro	3	2009	7,2	396	<Lq	6,8	0,013	80	5	3,8	nd	nd	11	nd	nd	nd	nd	nd	nd
Arraiolos	Santana do Campo	7	2009	7,7	483	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Santana do Campo	2	2007	7,5	463	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Santana do Campo	4	2007	7,8	500	<Lq	5	<Lq	<Lq	0,65	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Arraiolos	Santana do Campo	9	2007	7,6	820	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Santana do Campo	10	2007	7,6	690	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Santana do Campo	4	2008	7,8	520	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Santana do Campo	7	2008	7,7	790	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Santana do Campo	4	2009	7,6	710	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vale do Pereiro	9	2008	7,3	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vale do Pereiro	2	2009	7,3	710	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vale do Pereiro	6	2009	7,4	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vale do Pereiro	8	2007	7,3	750	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vale do Pereiro	10	2007	7,3	730	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vale do Pereiro	4	2007	7,3	740	<Lq	14	<Lq	<Lq	0,26	<Lq	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vale do Pereiro	1	2007	7,4	750	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vimieiro	6	2007	7,8	860	<Lq	41	<Lq	<Lq	<Lq	<Lq	nd	nd	6,5	nd	nd	nd	nd	nd	nd
Arraiolos	Vimieiro	11	2008	7,8	900	<Lq	41	<Lq	<Lq	<Lq	<Lq	4,1	80	5,2	<Lq	<Lq	0,2	<Lq	6	nd
Arraiolos	Vimieiro	11	2009	7,7	600	<Lq	14,5	<Lq	<Lq	<Lq	<Lq	2,4	50	<Lq	<Lq	<Lq	0,03	<Lq	<Lq	nd
Arraiolos	Vimieiro	2	2009	7,8	880	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vimieiro	1	2007	7,9	800	<Lq	38	<Lq	<Lq	<Lq	1,01	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Arraiolos	Vimieiro	7	2007	8	870	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vimieiro	3	2007	7,9	800	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vimieiro	3	2007	7,9	800	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vimieiro	10	2007	8	890	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vimieiro	11	2007	8	880	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vimieiro	3	2008	8	900	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vimieiro	5	2008	8	900	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vimieiro	6	2009	7,8	880	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Arraiolos	Vimieiro	5	2007	7,9	830	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Aldeia de Nora	2	2007	7,7	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Aldeia de Nora	4	2007	7,7	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Aldeia de Nora	6	2007	7,7	620	<Lq	17	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Borba	Aldeia de Nora	4	2008	7,5	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Aldeia de Nora	7	2008	7,8	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Aldeia de Nora	4	2009	7,3	660	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Aldeia de Nora	10	2009	7,4	660	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Aldeia de Sande	4	2007	7,7	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Aldeia de Sande	6	2007	7,5	630	<Lq	22	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Borba	Aldeia de Sande	4	2008	7,7	620	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Aldeia de Sande	7	2008	7,4	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Aldeia de Sande	4	2009	7,5	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Aldeia de Sande	10	2009	7,4	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Aldeia de Sande	2	2007	7,5	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Barro Branco	6	2007	7,5	640	<Lq	17	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Borba	Barro Branco	4	2008	7,5	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Barro Branco	7	2008	7,5	650	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Barro Branco	4	2009	7,2	660	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Barro Branco	10	2009	7,4	660	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Barro Branco	2	2007	7,5	660	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Barro Branco	4	2007	7,5	650	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Borba	4	2008	7,5	630	<Lq	24	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Borba	Borba	9	2008	7,6	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Borba	3	2007	7,9	630	<Lq	22	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Borba	Borba	5	2007	7,8	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd

Concelho	Localidade/Freguesia	mês	Ano	pH	Condutividade	Amónio	Nitrato	Nitrito	Ferro	Turvação	Oxidabilidade	Potássio	Sódio	Manganês	Cádmio	Crómio	Cobre	Níquel	Chumbo	Alumínio
Borba	Borba	6	2007	7,9	630	<Lq	24	<Lq	<Lq	<Lq	<Lq	0,9	9,2	<Lq	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Borba	Borba	6	2009	7,5	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Borba	1	2007	7,6	630	<Lq	23	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Borba	Borba	6	2008	7,5	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Borba	2	2009	7,2	630	<Lq	24	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Borba	Borba	10	2009	7,4	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Borba	2	2007	7,7	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Orada	2	2007	7,8	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Orada	4	2007	7,8	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Orada	6	2007	7,8	620	<Lq	21	<Lq	<Lq	<Lq	<Lq	nd	nd	9,3	nd	nd	nd	nd	nd	nd
Borba	Orada	4	2008	7,7	620	<Lq	23	<Lq	<Lq	<Lq	<Lq	0,76	9	<Lq	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Borba	Orada	9	2008	7,8	620	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Orada	2	2009	7,3	620	<Lq	24	<Lq	<Lq	<Lq	<Lq	0,8	9	<Lq	<Lq	<Lq	0,03	<Lq	<Lq	nd
Borba	Orada	7	2009	7,4	610	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Rio de Moinhos	1	2007	7,5	670	<Lq	16	<Lq	<Lq	0,31	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Borba	Rio de Moinhos	2	2007	7,6	660	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Rio de Moinhos	3	2007	7,7	650	<Lq	16	<Lq	<Lq	0,34	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Borba	Rio de Moinhos	5	2007	7,4	650	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Rio de Moinhos	6	2007	7,6	650	<Lq	19	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Borba	Rio de Moinhos	4	2008	7,5	650	<Lq	19	<Lq	<Lq	<Lq	<Lq	0,42	10	<Lq	<Lq	<Lq	0,022	<Lq	<Lq	nd
Borba	Rio de Moinhos	7	2009	7,2	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Rio de Moinhos	9	2008	7,8	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Borba	Rio de Moinhos	2	2009	7,3	660	<Lq	21	<Lq	<Lq	<Lq	<Lq	0,5	9	<Lq	<Lq	<Lq	0,01	<Lq	<Lq	nd
Estremoz	Arcos	2	2009	7,4	670	<Lq	45	<Lq	<Lq	0,8	<Lq	4,3	14	<Lq	<Lq	<Lq	0,04	<Lq	<Lq	nd
Estremoz	Arcos	1	2008	7,6	690	<Lq	29	<Lq	<Lq	<Lq	<Lq	1,9	12	<Lq	<Lq	<Lq	0,011	<Lq	<Lq	nd
Estremoz	Arcos	7	2008	7,5	690	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Arcos	11	2008	7,4	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Arcos	7	2009	7,3	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Arcos	10	2009	7,4	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Arcos	3	2007	7,3	660	<Lq	59	<Lq	50	0,64	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	Arcos	6	2007	7,4	660	<Lq	43	<Lq	<Lq	1	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	Espinheiro	3	2008	7,8	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Espinheiro	3	2009	7,8	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Espinheiro	10	2009	7,7	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Espinheiro	8	2008	7,9	720	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Estremoz	1	2007	7,7	730	<Lq	31	<Lq	<Lq	0,29	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	Estremoz	1	2008	7,7	720	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Estremoz	4	2008	7,8	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Estremoz	10	2008	7,6	720	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Estremoz	2	2009	7,6	650	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Estremoz	4	2009	7,4	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Estremoz	8	2009	7,6	710	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Estremoz	3	2007	7,6	740	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Estremoz	9	2007	7,8	720	<Lq	26	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	Estremoz	5	2007	7,7	720	<Lq	23	<Lq	<Lq	0,34	<Lq	2,3	18	<Lq	<Lq	<Lq	0,008	<Lq	<Lq	nd
Estremoz	Évoramonte	2	2009	7,9	710	<Lq	15,6	<Lq	<Lq	<Lq	<Lq	3	50	<Lq	<Lq	<Lq	0,007	<Lq	<Lq	nd
Estremoz	Évoramonte	1	2007	7,4	1040	<Lq	31	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	Évoramonte	1	2008	7,7	550	<Lq	2,9	<Lq	<Lq	<Lq	<Lq	1,4	42	<Lq	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Estremoz	Évoramonte	7	2008	7,4	1060	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Évoramonte	7	2009	7,2	550	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Évoramonte	3	2007	7,4	760	<Lq	15	<Lq	50	0,46	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	Évoramonte	6	2007	7,4	730	<Lq	16	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	Glória	1	2008	7,5	680	<Lq	35	<Lq	<Lq	<Lq	<Lq	2	9	<Lq	<Lq	<Lq	0,013	<Lq	<Lq	nd
Estremoz	Glória	7	2008	7,7	670	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Glória	2	2009	7,6	670	<Lq	48	<Lq	<Lq	<Lq	<Lq	2,1	14	<Lq	<Lq	<Lq	0,01	<Lq	<Lq	nd
Estremoz	Glória	3	2007	7,3	680	<Lq	38	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	Glória	6	2007	7,3	680	<Lq	36	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd

Concelho	Localidade/Freguesia	mês	Ano	pH	Condutividade	Amónio	Nitrato	Nitrito	Ferro	Turvação	Oxidabilidade	Potássio	Sódio	Manganês	Cádmio	Crómio	Cobre	Níquel	Chumbo	Alumínio
Estremoz	São Bento de Ana Loura	3	2008	7,5	480	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	São Bento de Ana Loura	9	2008	7,3	480	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	São Bento de Ana Loura	3	2009	7,1	464	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	São Bento de Ana Loura	3	2008	6,7	282	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	São Bento de Ana Loura	9	2008	7	252	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	São Bento do Ameixial	5	2008	7,7	680	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	São Bento do Ameixial	9	2007	7,6	460	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	São Bento do Ameixial	10	2007	7,7	450	<Lq	24	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	São Bento do Ameixial	10	2008	6,6	590	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	São Bento do Ameixial	5	2009	7	488	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	São Bento do Ameixial	12	2009	7,4	451	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	São Bento do Cortiço	3	2008	7,6	720	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	São Bento do Cortiço	9	2008	7,6	730	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	São Bento do Cortiço	3	2009	7,3	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	São Bento do Cortiço	11	2009	7,7	720	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	São Bento do Cortiço	9	2007	7,6	730	<Lq	22	<Lq	<Lq	0,32	<Lq	nd	nd	5	nd	nd	nd	nd	nd	nd
Estremoz	São Bento do Cortiço	1	2007	7,6	730	<Lq	29	<Lq	<Lq	0,4	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	S. Lourenço de Mamporcão	3	2009	7,2	750	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	S. Lourenço de Mamporcão	1	2007	7,5	770	<Lq	36	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	S. Lourenço de Mamporcão	3	2008	7,5	760	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	S. Lourenço de Mamporcão	9	2008	7,6	760	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	S. Lourenço de Mamporcão	9	2007	7,4	760	<Lq	37	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	S. Lourenço de Mamporcão	5	2007	7,5	760	<Lq	34	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	Santo Estêvão	3	2008	7,8	560	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	São Bento de Ana Loura	1	2007	7,2	490	<Lq	2,8	<Lq	<Lq	0,32	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	São Bento de Ana Loura	9	2007	7,2	490	<Lq	2,9	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	São Bento de Ana Loura	1	2007	6,8	255	<Lq	7,9	<Lq	<Lq	<Lq	<Lq	nd	nd	17	nd	nd	nd	nd	nd	nd
Estremoz	São Bento de Ana Loura	9	2007	6,7	277	<Lq	6,1	<Lq	<Lq	<Lq	<Lq	nd	nd	23	nd	nd	nd	nd	nd	nd
Estremoz	S. Domingos de Ana Loura	3	2008	7,7	710	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	S. Domingos de Ana Loura	10	2009	7,6	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	S. Domingos de Ana Loura	2	2007	7,5	680	<Lq	47	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	S. Domingos de Ana Loura	8	2008	8	720	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	S. Domingos de Ana Loura	3	2009	7,7	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	S. Domingos de Ana Loura	2	2007	7,6	700	<Lq	40	<Lq	<Lq	0,51	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	Sta Vitória do Ameixial	9	2007	7,5	700	<Lq	18	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	Sta Vitória do Ameixial	12	2007	7,7	700	<Lq	19	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	Sta Vitória do Ameixial	5	2008	7,7	690	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Sta Vitória do Ameixial	10	2008	7,6	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Sta Vitória do Ameixial	5	2009	7,4	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Sta Vitória do Ameixial	12	2009	7,5	590	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Santo Estêvão	9	2008	7,8	560	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Santo Estêvão	3	2009	7,1	570	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Santo Estêvão	1	2007	7,5	580	<Lq	13	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	Santo Estêvão	9	2007	7,8	560	<Lq	11,4	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	Veiros	9	2007	7,6	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Veiros	12	2007	7,8	710	<Lq	43	<Lq	<Lq	<Lq	1,4	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Estremoz	Veiros	5	2008	8	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Veiros	10	2008	7,6	680	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Veiros	2	2009	7,5	463	<Lq	32	<Lq	<Lq	3,7	<Lq	3,6	26	<Lq	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Estremoz	Veiros	5	2009	7,6	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Veiros	2	2008	7,7	690	<Lq	51	<Lq	<Lq	<Lq	<Lq	4,8	37	6,5	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Estremoz	Veiros	12	2009	7,8	520	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Veiros	5	2007	7,6	710	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Venda da Porca	5	2008	7,9	570	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Venda da Porca	10	2008	7,8	560	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Venda da Porca	5	2009	7,6	580	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Estremoz	Venda da Porca	12	2009	7,7	570	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd

Concelho	Localidade/Freguesia	mês	Ano	pH	Condutividade	Amónio	Nitrato	Nitrato	Ferro	Turvação	Oxidabilidade	Potássio	Sódio	Manganês	Cádmio	Crómio	Cobre	Níquel	Chumbo	Alumínio
Estremoz	Venda da Porca	5	2007	7,8	580	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Horta das Figueiras	5	2007	7,9	359	<Lq	7	<Lq	<Lq	0,26	2	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	Horta das Figueiras	6	2007	7,7	381	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Horta das Figueiras	8	2007	7,6	408	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Horta das Figueiras	9	2007	7,6	404	<Lq	3,5	<Lq	<Lq	0,28	1,9	nd	nd	23	nd	nd	nd	nd	nd	nd
Évora	Horta das Figueiras	10	2007	7,6	410	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Horta das Figueiras	12	2007	7,7	410	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Horta das Figueiras	5	2008	7,6	403	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Horta das Figueiras	9	2008	7,6	434	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Azaruja	9	2009	7,6	428	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Azaruja	9	2009	7,6	428	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Azaruja	10	2007	7,7	600	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Azaruja	2	2008	7,8	610	<Lq	23	<Lq	<Lq	0,29	<Lq	2,6	49	<Lq	<Lq	<Lq	0,013	<Lq	<Lq	nd
Évora	Azaruja	2	2007	7,6	354	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Azaruja	6	2007	7,8	590	<Lq	22	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	Azaruja	9	2008	7,6	600	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Azaruja	2	2009	7,6	560	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Boa Fé	3	2008	7,7	374	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Boa Fé	3	2009	7,6	385	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Boa Fé	9	2009	7,2	406	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Boa Fé	1	2007	7,8	377	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Boa Fé	4	2007	7,6	374	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Boa Fé	10	2007	7,8	381	<Lq	<Lq	<Lq	<Lq	9	<Lq	nd	nd	86	nd	nd	nd	nd	nd	nd
Évora	Boa Fé	10	2008	7,4	381	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Carvalhas	4	2007	7,8	372	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Carvalhas	10	2007	7,4	390	<Lq	2,1	<Lq	<Lq	5,1	<Lq	nd	nd	48	nd	nd	nd	nd	nd	nd
Évora	Carvalhas	3	2008	7,4	390	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Carvalhas	10	2008	7,4	387	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Carvalhas	3	2009	7,6	386	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Carvalhas	9	2009	7,7	406	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Casas Novas	4	2007	7,8	373	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Casas Novas	10	2007	7,8	381	<Lq	2,1	<Lq	<Lq	1,9	<Lq	nd	nd	31	nd	nd	nd	nd	nd	nd
Évora	Casas Novas	3	2008	7,8	379	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Casas Novas	10	2008	7,4	384	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Casas Novas	3	2009	7,3	386	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Casas Novas	9	2009	7,3	405	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Alcáçovas	8	2007	7,6	690	<Lq	12	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	Alcáçovas	4	2007	7,4	540	<Lq	24	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	Alcáçovas	2	2007	7,4	860	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Alcáçovas	3	2007	7,4	860	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Alcáçovas	5	2008	8,4	427	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Alcáçovas	9	2008	8	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Alcáçovas	1	2009	7,6	530	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Alcáçovas	9	2009	7,7	510	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Sé e São Pedro	6	2009	7,3	452	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	110
Évora	Sé e São Pedro	1	2007	7,6	340	<Lq	13	<Lq	<Lq	<Lq	1,34	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	Sé e São Pedro	2	2007	7,9	354	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Sé e São Pedro	3	2007	7,8	349	<Lq	10,5	<Lq	<Lq	<Lq	1,6	nd	nd	5,9	nd	nd	nd	nd	nd	nd
Évora	Sé e São Pedro	4	2007	7,9	369	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Sé e São Pedro	7	2007	7,8	392	<Lq	3,9	<Lq	<Lq	0,3	2,1	nd	nd	11,2	nd	nd	nd	nd	nd	nd
Évora	Sé e São Pedro	8	2007	7,7	407	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Sé e São Pedro	9	2007	7,5	404	<Lq	3,5	<Lq	<Lq	0,27	1,9	nd	nd	30	nd	nd	nd	nd	nd	nd
Évora	Sé e São Pedro	10	2007	7,7	410	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Sé e São Pedro	11	2007	7,7	401	<Lq	3,7	<Lq	<Lq	<Lq	2	nd	nd	13,4	nd	nd	nd	nd	nd	nd
Évora	Sé e São Pedro	12	2007	7,7	410	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Sé e São Pedro	1	2008	7,8	420	<Lq	4,3	<Lq	<Lq	<Lq	1,6	3,8	35	8,9	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Évora	Sé e São Pedro	6	2008	7,6	401	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd

Concelho	Localidade/Freguesia	mês	Ano	pH	Condutividade	Amónio	Nitrato	Nitrito	Ferro	Turvação	Oxidabilidade	Potássio	Sódio	Manganês	Cádmio	Crómio	Cobre	Níquel	Chumbo	Alumínio
Évora	Sé e São Pedro	11	2008	7,5	403	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	<Lq
Évora	Horta das Figueiras	1	2007	7,6	342	<Lq	13	<Lq	<Lq	<Lq	1,28	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	Horta das Figueiras	3	2007	7,9	352	<Lq	10,5	<Lq	<Lq	<Lq	1,4	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	Horta das Figueiras	4	2007	8	371	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Horta das Figueiras	11	2007	7,7	401	<Lq	3,6	<Lq	<Lq	0,27	1,9	3,8	36	10,1	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Évora	Malagueira	1	2007	7,6	342	<Lq	12	<Lq	<Lq	<Lq	1,35	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	Malagueira	2	2007	7,8	354	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Malagueira	3	2007	7,8	351	<Lq	10,7	<Lq	<Lq	0,37	1,27	nd	nd	10,3	nd	nd	nd	nd	nd	nd
Évora	Malagueira	4	2007	8	372	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Malagueira	5	2007	7,9	360	<Lq	7,4	<Lq	<Lq	<Lq	2,1	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	Malagueira	6	2007	7,7	384	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Malagueira	7	2007	7,8	394	<Lq	4	<Lq	<Lq	0,5	2,2	nd	nd	17	nd	nd	nd	nd	nd	nd
Évora	Malagueira	8	2007	7,6	410	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Malagueira	9	2007	7,6	405	<Lq	3,7	<Lq	<Lq	0,25	1,7	nd	nd	23	nd	nd	nd	nd	nd	nd
Évora	Malagueira	10	2007	7,6	410	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Malagueira	11	2007	7,7	400	<Lq	3,7	<Lq	<Lq	<Lq	2,2	nd	nd	10,4	nd	nd	nd	nd	nd	nd
Évora	Malagueira	12	2007	7,8	410	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Malagueira	5	2008	7,7	378	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Malagueira	9	2008	7,5	434	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Bacêlo	1	2007	7,5	343	<Lq	12	<Lq	<Lq	<Lq	1,36	nd	nd	<Lq	nd	nd	nd	nd	nd	53
Évora	Bacêlo	2	2007	7,8	340	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	<Lq
Évora	Bacêlo	3	2007	7,8	326	<Lq	10,9	<Lq	<Lq	<Lq	1,6	nd	nd	6,5	nd	nd	nd	nd	nd	47
Évora	Bacêlo	4	2007	8	369	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Bacêlo	5	2007	8	359	<Lq	7,7	<Lq	<Lq	<Lq	1,8	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Bacêlo	6	2007	7,7	384	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Bacêlo	7	2007	7,8	393	<Lq	3,9	<Lq	<Lq	<Lq	2,1	nd	nd	9,4	nd	nd	nd	nd	nd	32
Évora	Bacêlo	8	2007	7,6	408	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	52
Évora	Bacêlo	9	2007	7,6	406	<Lq	3,5	<Lq	<Lq	0,28	2	nd	nd	30	nd	nd	nd	nd	nd	<Lq
Évora	Bacêlo	10	2007	7,8	410	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	35
Évora	Bacêlo	11	2007	7,7	400	<Lq	3,7	<Lq	<Lq	<Lq	2,2	nd	nd	15	nd	nd	nd	nd	nd	36
Évora	Bacêlo	12	2007	7,7	410	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	39
Évora	Bacêlo	5	2008	7,6	404	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Bacêlo	9	2008	7,6	435	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Bacêlo	2	2009	7,2	356	<Lq	4,8	<Lq	<Lq	0,7	2,7	3,8	25	17	<Lq	<Lq	<Lq	<Lq	<Lq	160
Évora	Bacêlo	8	2009	7,4	422	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	110
Évora	Horta das Figueiras	2	2007	8	357	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Sé e São Pedro	5	2007	8	361	<Lq	6,8	<Lq	<Lq	1,25	1,8	nd	nd	8,6	nd	nd	nd	nd	nd	nd
Évora	Sé e São Pedro	6	2007	7,7	384	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Sé e São Pedro	10	2009	7,3	444	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	130
Évora	Horta das Figueiras	7	2007	7,7	392	<Lq	3,9	<Lq	<Lq	1,1	2,1	nd	nd	10,1	nd	nd	nd	nd	nd	nd
Évora	Horta das Figueiras	12	2009	7,7	449	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	110
Évora	Malagueira	4	2009	7,3	400	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	90
Évora	Nª Sra Graça do Divor	10	2007	7,3	375	<Lq	4	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	Nª Sra Graça do Divor	3	2008	7,1	362	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Nª Sra Graça do Divor	10	2008	7,3	388	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Nª Sra Graça do Divor	3	2009	7,1	364	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Nª Sra Graça do Divor	4	2007	7,5	364	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Nª Sra Graça do Divor	9	2009	7	335	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Nª Sra de Guadalupe	2	2007	7,7	344	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Nª Sra de Guadalupe	4	2007	8	369	<Lq	8,5	<Lq	<Lq	0,38	1,7	nd	nd	14	nd	nd	nd	nd	nd	nd
Évora	Nª Sra de Guadalupe	8	2007	7,9	411	<Lq	3,4	<Lq	<Lq	0,49	2,1	nd	nd	18	nd	nd	nd	nd	nd	nd
Évora	Nª Sra de Guadalupe	5	2008	7,8	407	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Nª Sra de Guadalupe	9	2008	7,9	441	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Nª Sra de Guadalupe	1	2009	7,9	459	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Sé e São Pedro	4	2009	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	140
Évora	Sé e São Pedro	6	2009	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	88
Évora	Sé e São Pedro	4	2009	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	140

Concelho	Localidade/Freguesia	mês	Ano	pH	Condutividade	Amónio	Nitrato	Nitrato	Ferro	Turvação	Oxidabilidade	Potássio	Sódio	Manganês	Cádmio	Crómio	Cobre	Níquel	Chumbo	Alumínio
Évora	Nº Sra de Guadalupe	3	2007	7,6	354	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Nº Sra de Machede	6	2007	8	401	<Lq	<Lq	<Lq	<Lq	<Lq	1,9	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	Nº Sra de Machede	10	2007	7,8	430	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Nº Sra de Machede	2	2008	7,9	440	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Nº Sra de Machede	9	2008	7,9	439	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Nº Sra de Machede	2	2009	7,7	450	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Nº Sra de Machede	2	2007	8,2	390	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Nº Sra Graça do Divor	1	2007	7,2	368	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Nº Sra de Machede	9	2009	7,5	433	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Bacêlo	4	2009	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	130
Évora	Bacêlo	4	2009	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	130
Évora	Bacêlo	6	2009	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	65
Évora	São Bento	4	2009	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	140
Évora	São Bento	6	2009	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	150
Évora	São Bento	4	2009	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	170
Évora	São Bento	4	2009	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	170
Évora	São Bento	6	2009	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	100
Évora	São Bento	4	2009	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	140
Évora	S. Brás do Regedouro	4	2007	7,7	770	<Lq	14	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	S. Brás do Regedouro	8	2007	7,6	770	<Lq	13	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	S. Brás do Regedouro	5	2008	7,7	760	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Brás do Regedouro	9	2008	7,7	760	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Brás do Regedouro	1	2009	7,7	760	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Brás do Regedouro	9	2009	7,4	760	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Brás do Regedouro	2	2007	7,6	790	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	São Manços	3	2007	7,5	730	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	São Manços	7	2007	7,3	720	<Lq	37	0,018	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	São Manços	11	2007	7,3	740	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	São Manços	1	2008	7,3	730	<Lq	42	<Lq	<Lq	<Lq	<Lq	2,7	42	<Lq	<Lq	<Lq	0,014	<Lq	<Lq	nd
Évora	São Manços	9	2008	7,5	429	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	São Manços	3	2009	7,2	371	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	São Manços	9	2009	7,6	446	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Miguel de Machede	6	2007	7,7	580	<Lq	40	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	S. Miguel de Machede	10	2007	7,5	570	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Miguel de Machede	2	2008	7,7	580	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Miguel de Machede	9	2008	7,5	620	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Miguel de Machede	2	2009	7,5	600	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Miguel de Machede	2	2007	7,6	570	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Miguel de Machede	9	2009	7,5	431	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Seb. da Giesteira	4	2007	7,7	375	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Seb. da Giesteira	10	2007	7,7	430	<Lq	<Lq	<Lq	<Lq	<Lq	2,3	nd	nd	5,4	nd	nd	nd	nd	nd	nd
Évora	S. Seb. da Giesteira	3	2008	7,6	411	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Seb. da Giesteira	10	2008	7,3	470	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Seb. da Giesteira	3	2009	7,5	365	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Seb. da Giesteira	1	2007	7,8	353	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Seb. da Giesteira	9	2009	7	468	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Vicente de Valongo	3	2007	7,7	455	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Vicente de Valongo	7	2007	7,8	434	<Lq	36	<Lq	<Lq	<Lq	1,14	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	S. Vicente de Valongo	10	2008	7,9	460	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Vicente de Valongo	9	2009	7,3	425	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Vicente de Valongo	11	2007	8,1	440	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	S. Vicente de Valongo	1	2008	8	440	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Torre de Coelheiros	3	2007	7,4	800	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Torre de Coelheiros	7	2007	7,4	870	<Lq	30	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	Torre de Coelheiros	11	2007	7,4	840	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Torre de Coelheiros	1	2008	7,4	850	<Lq	37	<Lq	<Lq	<Lq	<Lq	4,6	80	<Lq	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Évora	Torre de Coelheiros	9	2008	7,4	850	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd

Concelho	Localidade/Freguesia	mês	Ano	pH	Condutividade	Amónio	Nitrato	Nitrito	Ferro	Turvação	Oxidabilidade	Potássio	Sódio	Manganês	Cádmio	Crómio	Cobre	Níquel	Chumbo	Alumínio
Évora	Torre de Coelheiros	3	2009	7,2	850	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Torre de Coelheiros	9	2009	7,4	970	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Valongo	3	2009	7,5	378	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Valverde	2	2007	7,9	351	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Valverde	4	2007	7,1	378	<Lq	56	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Évora	Valverde	8	2007	7,8	410	<Lq	3,7	<Lq	<Lq	<Lq	2,1	nd	nd	9,7	nd	nd	nd	nd	nd	nd
Évora	Valverde	1	2008	7,9	420	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Valverde	5	2008	7,9	411	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Valverde	9	2008	7,8	447	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Valverde	1	2009	8	459	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Valverde	9	2009	7,2	491	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Valverde	10	2009	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Vendinha	3	2007	7,7	670	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Vendinha	7	2007	7,6	680	<Lq	42	<Lq	<Lq	<Lq	<Lq	nd	nd	7,5	nd	nd	nd	nd	nd	nd
Évora	Vendinha	11	2007	7,6	730	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Vendinha	1	2008	7,6	810	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Vendinha	10	2008	7,7	451	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Vendinha	3	2009	7,3	375	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Évora	Vendinha	9	2009	7,4	430	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mora	Brotas	4	2007	6,5	184	<Lq	<Lq	<Lq	<Lq	0,32	<Lq	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mora	Cabeção	5	2007	7,4	322	<Lq	2,5	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Mora	Cabeção	5	2007	7,5	327	<Lq	2,6	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Mora	Cabeção	2	2009	7,4	327	<Lq	2,5	<Lq	<Lq	0,7	<Lq	1,5	37	<Lq	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Mora	Malarranha	5	2007	7,5	910	<Lq	2,4	<Lq	<Lq	<Lq	1,05	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Mora	Malarranha	6	2008	nd	nd	nd	nd	nd	1500	nd	nd	nd	nd	88	nd	nd	nd	nd	nd	nd
Mora	Mora	4	2007	6,3	152	<Lq	1,9	<Lq	<Lq	<Lq	<Lq	1,4	25	<Lq	<Lq	<Lq	0,02	<Lq	<Lq	nd
Mora	Mora	6	2007	6,4	202	<Lq	18	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Mora	Mora	3	2007	7,2	201	<Lq	19	<Lq	<Lq	<Lq	<Lq	2,4	24	<Lq	<Lq	<Lq	0,016	<Lq	<Lq	nd
Mora	Mora	4	2007	6,3	152	<Lq	1,9	<Lq	<Lq	<Lq	<Lq	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mora	Pavia	5	2007	7,4	329	<Lq	2,5	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Mourão	Granja	4	2008	7,4	870	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mourão	Granja	8	2008	7,3	920	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mourão	Granja	4	2009	7,4	890	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mourão	Granja	10	2009	7,1	1000	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mourão	Granja	1	2007	7,4	750	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mourão	Granja	4	2007	7,6	750	<Lq	1,6	<Lq	<Lq	<Lq	<Lq	nd	nd	8,8	nd	nd	nd	nd	nd	nd
Mourão	Luz	1	2007	7,4	900	<Lq	39	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Mourão	Luz	3	2007	7,2	690	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mourão	Luz	10	2007	7,4	890	<Lq	40	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Mourão	Luz	4	2008	7,5	840	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mourão	Luz	8	2008	7,4	920	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mourão	Luz	4	2009	7,4	950	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mourão	Luz	10	2009	7,3	950	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mourão	Mourão	2	2007	7,4	710	<Lq	36	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Mourão	Mourão	5	2007	7,5	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mourão	Mourão	9	2007	7,5	630	<Lq	27	<Lq	<Lq	<Lq	<Lq	0,41	38	<Lq	<Lq	<Lq	0,011	<Lq	<Lq	nd
Mourão	Mourão	1	2008	7,5	640	<Lq	31	<Lq	<Lq	<Lq	<Lq	0,42	35	<Lq	<Lq	<Lq	0,009	<Lq	<Lq	nd
Mourão	Mourão	4	2008	7,6	760	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mourão	Mourão	8	2008	7,5	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mourão	Mourão	1	2009	7,8	467	<Lq	2,7	<Lq	<Lq	<Lq	2,7	3,9	41	8,9	<Lq	<Lq	0,01	<Lq	<Lq	nd
Mourão	Mourão	4	2009	7,7	378	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Mourão	Mourão	10	2009	7,6	690	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Alqueva	7	2009	7,6	477	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Alqueva	3	2009	7,7	404	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Alqueva	2	2007	7,4	374	<Lq	2,9	<Lq	<Lq	<Lq	2,2	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Portel	Alqueva	3	2007	7,6	377	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Alqueva	7	2007	7,5	500	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd

Concelho	Localidade/Freguesia	mês	Ano	pH	Condutividade	Amónio	Nitrato	Nitrito	Ferro	Turvação	Oxidabilidade	Potássio	Sódio	Manganês	Cádmio	Crómio	Cobre	Níquel	Chumbo	Alumínio
Portel	Alqueva	10	2007	7,6	520	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Alqueva	3	2008	7,4	480	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Alqueva	7	2008	7,6	427	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Amieira	3	2009	7,6	403	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Amieira	7	2009	7,4	620	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Amieira	3	2007	6,8	159	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Amieira	3	2007	6,6	147	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Amieira	3	2007	6,8	149	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Amieira	7	2008	7,7	432	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Amieira	2	2007	7,5	382	<Lq	3	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Portel	Amieira	3	2007	7,4	365	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Amieira	3	2008	7,5	480	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Amieira	10	2007	7,8	530	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Amieira	7	2007	7,3	460	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Monte do Trigo	3	2009	7,4	403	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Monte do Trigo	2	2007	7,6	820	<Lq	49	<Lq	<Lq	<Lq	2,4	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Portel	Monte do Trigo	3	2007	7,3	820	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Monte do Trigo	7	2007	7,5	810	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Monte do Trigo	10	2007	7,5	800	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Monte do Trigo	3	2008	7,5	470	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Monte do Trigo	7	2009	7,6	690	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Monte do Trigo	11	2009	7,5	610	<Lq	7,8	<Lq	90	<Lq	<Lq	2,7	53	<Lq	<Lq	<Lq	0,008	<Lq	<Lq	nd
Portel	Monte do Trigo	7	2008	7,7	770	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Monte do Trigo	11	2008	7,6	780	<Lq	38	<Lq	<Lq	<Lq	<Lq	2	77	<Lq	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Portel	Oriola	11	2009	7,3	1060	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Oriola	2	2007	7,4	960	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Oriola	5	2007	7,2	1070	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Oriola	9	2007	7,3	1060	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Oriola	11	2007	7,4	1070	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Oriola	3	2008	7,4	1020	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Oriola	11	2008	7,4	1050	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Oriola	3	2009	7,3	1020	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Portel	3	2009	7,6	404	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Portel	7	2009	7	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Portel	1	2007	6,8	117	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Portel	1	2007	7,6	292	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Portel	1	2007	6,4	130	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Portel	1	2007	7	102	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Portel	1	2007	6,5	103	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Portel	1	2007	6,2	130	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Portel	1	2007	7,5	123	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Portel	1	2007	6,8	113	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Portel	10	2007	7,5	400	<Lq	1,7	<Lq	<Lq	<Lq	2,4	3,6	33	5,8	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Portel	Portel	7	2008	7,6	423	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Portel	3	2007	7,2	357	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Portel	3	2008	7,6	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Portel	2	2007	7,5	364	<Lq	2,8	<Lq	<Lq	<Lq	2,7	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Portel	Portel	7	2007	7,5	384	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	S. B. do Outeiro	2	2007	7,3	391	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	S. B. do Outeiro	9	2007	7,3	800	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	S. B. do Outeiro	3	2008	7,3	430	<Lq	<Lq	<Lq	<Lq	0,29	1,6	nd	nd	13	nd	nd	nd	nd	nd	nd
Portel	S. B. do Outeiro	1	2007	7,2	790	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	S. B. do Outeiro	1	2007	7,1	466	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	S. B. do Outeiro	3	2009	7,5	404	<Lq	2,1	<Lq	<Lq	<Lq	1,7	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Portel	S. B. do Outeiro	9	2009	7,5	530	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	S. B. do Outeiro	9	2008	7,5	439	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	S. B. do Outeiro	5	2007	7,3	379	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd

Concelho	Localidade/Freguesia	mês	Ano	pH	Condutividade	Amónio	Nitrato	Nitrato	Ferro	Turvação	Oxidabilidade	Potássio	Sódio	Manganês	Cádmio	Crómio	Cobre	Níquel	Chumbo	Alumínio
Portel	S. B. do Outeiro	11	2007	7,6	820	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Santana	3	2009	7,6	444	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Santana	9	2009	7,6	590	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Santana	9	2008	7,9	129	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Santana	2	2007	7,3	660	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Santana	5	2007	7,3	510	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Santana	9	2007	7,3	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Santana	11	2007	7,5	680	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Santana	3	2008	7,5	560	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Vera Cruz	9	2008	7,6	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Vera Cruz	3	2009	7,5	710	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Vera Cruz	9	2009	7,6	660	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Vera Cruz	4	2007	7,4	231	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Vera Cruz	4	2007	7,5	900	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Vera Cruz	4	2007	7,4	830	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Vera Cruz	3	2008	7,6	650	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Vera Cruz	9	2007	7,6	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Vera Cruz	11	2007	7,8	660	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Vera Cruz	2	2007	7,5	680	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Portel	Vera Cruz	5	2007	7,5	670	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Aldeia da Serra	1	2007	7,2	433	<Lq	10	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Redondo	Aldeia da Serra	5	2007	6,7	470	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Aldeia da Serra	4	2008	7,2	530	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Aldeia da Serra	10	2008	7,6	480	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Aldeia da Serra	4	2009	7,2	374	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Aldeias de Montoito	2	2007	7,4	332	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Aldeias de Montoito	6	2007	7,4	346	<Lq	<Lq	<Lq	<Lq	0,28	2,8	nd	nd	37	nd	nd	nd	nd	nd	nd
Redondo	Aldeias de Montoito	10	2007	7,6	353	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Aldeias de Montoito	3	2008	7,3	403	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Aldeias de Montoito	9	2008	7,4	357	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Aldeias de Montoito	6	2009	7,7	346	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Falcoeirias	1	2007	7,3	322	<Lq	5,9	<Lq	<Lq	<Lq	3,4	nd	nd	9,3	nd	nd	nd	nd	nd	nd
Redondo	Falcoeirias	6	2007	7,4	338	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Falcoeirias	10	2007	6,8	327	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Falcoeirias	4	2008	7,6	316	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Falcoeirias	8	2008	7,8	358	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Falcoeirias	9	2009	7,8	396	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Foros Fonte Seca	2	2007	7,6	346	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Foros Fonte Seca	10	2007	7,7	369	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Foros Fonte Seca	3	2008	7,8	426	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Foros Fonte Seca	9	2008	8,4	375	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Foros Fonte Seca	3	2009	8,4	326	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Foros da Fontana	2	2007	7,5	365	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Foros da Fontana	5	2007	7,6	351	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Foros da Fontana	8	2007	7,5	385	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Foros da Fontana	4	2008	7,9	340	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Foros da Fontana	8	2008	7,9	368	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Foros da Fontana	5	2009	7,5	350	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Montoito	2	2007	7,4	332	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Montoito	6	2007	7,4	344	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Montoito	10	2007	7,6	353	<Lq	1,8	<Lq	<Lq	<Lq	2,3	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Redondo	Montoito	3	2008	7,5	406	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	410
Redondo	Montoito	9	2008	8,2	365	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	370
Redondo	Montoito	3	2009	7,6	357	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	41
Redondo	Redondo	1	2007	8,1	357	<Lq	5,7	<Lq	<Lq	0,42	2,8	nd	nd	5,8	nd	nd	nd	nd	nd	nd
Redondo	Redondo	2	2007	7,9	357	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Redondo	3	2007	8,4	343	<Lq	4,6	<Lq	<Lq	<Lq	1,8	nd	nd	5,3	nd	nd	nd	nd	nd	nd

Concelho	Localidade/Freguesia	mês	Ano	pH	Condutividade	Amónio	Nitrato	Nitrito	Ferro	Turvação	Oxidabilidade	Potássio	Sódio	Manganês	Cádmio	Crómio	Cobre	Níquel	Chumbo	Alumínio
Redondo	Redondo	4	2007	7,8	353	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Redondo	5	2007	8,4	353	<Lq	nd	<Lq	<Lq	1	2,5	nd	nd	15	nd	nd	nd	nd	nd	nd
Redondo	Redondo	6	2007	8,5	367	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Redondo	7	2007	8,4	378	<Lq	<Lq	<Lq	<Lq	<Lq	2,8	nd	nd	15	nd	nd	nd	nd	nd	nd
Redondo	Redondo	8	2007	8,1	388	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Redondo	9	2007	8,5	396	<Lq	1,9	<Lq	<Lq	<Lq	2,9	nd	nd	7,1	nd	nd	nd	nd	nd	nd
Redondo	Redondo	10	2007	8,7	369	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Redondo	11	2007	8,7	375	<Lq	1,9	<Lq	<Lq	<Lq	2,1	nd	nd	5,4	nd	nd	nd	nd	nd	nd
Redondo	Redondo	1	2008	8,4	450	<Lq	2	<Lq	<Lq	<Lq	2,4	4,1	35	5,9	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Redondo	Redondo	4	2008	8,4	349	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	133
Redondo	Redondo	9	2008	8,8	372	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	160
Redondo	Redondo	2	2009	7,8	378	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Redondo	10	2009	7,8	411	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Redondo	8	2007	7	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Redondo	11	2009	7,1	398	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Santa Susana	4	2008	7,5	322	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Santa Susana	4	2009	7,6	354	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Santa Susana	7	2007	7,3	344	<Lq	<Lq	<Lq	<Lq	0,26	2,4	nd	nd	42	nd	nd	nd	nd	nd	nd
Redondo	Santa Susana	1	2007	7,5	334	<Lq	5,7	<Lq	<Lq	0,3	2,6	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Redondo	Santa Susana	3	2007	7,2	323	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Santa Susana	11	2007	7,2	328	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Redondo	Santa Susana	10	2008	8	370	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Barrada	10	2009	7,8	445	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Barrada	7	2009	7,8	437	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Barrada	8	2008	7,7	600	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Barrada	1	2007	7,8	339	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Barrada	10	2008	8,3	430	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Campinho	7	2009	7,5	780	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Campinho	2	2008	7,6	900	<Lq	30	<Lq	<Lq	<Lq	<Lq	9	55	<Lq	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Reguengos de Monsaraz	Campinho	3	2007	7,7	890	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Campinho	2	2009	7,4	461	<Lq	9,1	<Lq	<Lq	<Lq	2	4,6	32	<Lq	<Lq	<Lq	0,01	<Lq	<Lq	nd
Reguengos de Monsaraz	Campinho	7	2008	7,5	850	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Caridade	6	2009	7,4	446	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Caridade	8	2008	7,8	432	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Caridade	10	2009	7,4	445	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Caridade	10	2008	8,1	460	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Caridade	1	2007	7,5	325	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Carrapatelo	9	2007	7,4	960	<Lq	78	<Lq	<Lq	0,51	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Carrapatelo	11	2008	7,6	460	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Carrapatelo	12	2008	7,8	470	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Carrapatelo	4	2009	7,5	377	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Carrapatelo	12	2009	7,7	457	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Cumeada	6	2009	6,8	393	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Cumeada	3	2007	7,2	362	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Cumeada	11	2007	7,4	344	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Cumeada	6	2008	7,8	405	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Cumeada	11	2009	7,7	446	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Cumeada	10	2008	8	460	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Ferragudo	1	2007	7,4	348	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Ferragudo	8	2008	7,8	560	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Ferragudo	6	2009	7,2	770	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Ferragudo	10	2008	7,7	428	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Ferragudo	10	2009	7,5	500	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Monsaraz	2	2007	7,6	332	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Monsaraz	6	2007	7,8	358	<Lq	2,4	<Lq	<Lq	0,3	2,4	nd	nd	12,7	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Monsaraz	2	2008	8	480	<Lq	9,5	0,014	<Lq	<Lq	1,4	3,8	37	5,1	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Reguengos de Monsaraz	Monsaraz	7	2008	7,8	500	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd

Concelho	Localidade/Freguesia	mês	Ano	pH	Condutividade	Amónio	Nitrato	Nitrato	Ferro	Turvação	Oxidabilidade	Potássio	Sódio	Manganês	Cádmio	Crómio	Cobre	Níquel	Chumbo	Alumínio
Reguengos de Monsaraz	Monsaraz	2	2009	7,5	374	<Lq	5	<Lq	<Lq	0,36	2,9	3,8	29	11	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Reguengos de Monsaraz	Monsaraz	7	2009	7,8	452	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Motrinos	4	2008	7,5	600	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Motrinos	9	2008	7,9	438	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Motrinos	4	2009	7,5	387	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Motrinos	9	2009	7,6	426	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Motrinos	2	2007	7,2	337	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Outeiro	5	2007	7,6	329	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Outeiro	4	2008	7,5	600	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Outeiro	4	2009	7,4	375	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Outeiro	9	2008	8	440	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Outeiro	10	2007	7,5	590	<Lq	22	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Outeiro	9	2009	7,9	431	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Perolivas	4	2007	7,5	321	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Perolivas	9	2007	7,5	367	<Lq	<Lq	<Lq	<Lq	<Lq	2,6	nd	nd	9,2	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Perolivas	8	2008	7,8	438	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Perolivas	10	2008	7,7	451	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Perolivas	6	2009	7,3	433	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Perolivas	10	2009	7,5	442	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Reguengos de Monsaraz	8	2007	7,4	346	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Reguengos de Monsaraz	10	2007	7,7	1220	<Lq	130	<Lq	<Lq	<Lq	<Lq	nd	nd	24	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Reguengos de Monsaraz	2	2008	7,9	420	<Lq	2	<Lq	<Lq	<Lq	1,8	nd	nd	9,8	nd	nd	nd	nd	nd	35
Reguengos de Monsaraz	Reguengos de Monsaraz	7	2008	7,6	434	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Reguengos de Monsaraz	12	2008	7,9	470	<Lq	1,8	<Lq	<Lq	<Lq	2,6	nd	nd	13	nd	nd	nd	nd	nd	170
Reguengos de Monsaraz	Reguengos de Monsaraz	2	2009	7,2	373	<Lq	4,5	<Lq	<Lq	0,28	2,8	nd	nd	12	nd	nd	nd	nd	nd	100
Reguengos de Monsaraz	Reguengos de Monsaraz	7	2009	7,7	444	<Lq	<Lq	<Lq	70	<Lq	2,5	nd	nd	20	nd	nd	nd	nd	nd	160
Reguengos de Monsaraz	Reguengos de Monsaraz	12	2009	7,6	444	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Reguengos de Monsaraz	5	2007	7,3	325	<Lq	nd	<Lq	<Lq	0,44	7,3	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Reguengos de Monsaraz	1	2007	7,5	341	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	R. Monsaraz	3	2007	7,5	348	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	S. Marcos do Campo	3	2007	7,1	345	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	S. Marcos do Campo	6	2008	8	407	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	S. Marcos do Campo	10	2008	8	460	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	S. Marcos do Campo	6	2009	7	424	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	S. Marcos do Campo	11	2009	7,7	445	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	S. Pedro do Corval	6	2009	7,4	428	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	S. Pedro do Corval	11	2009	7	438	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	S. Pedro do Corval	6	2007	7,5	342	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	S. Pedro do Corval	11	2007	7,6	400	<Lq	13	<Lq	<Lq	<Lq	2	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	S. Pedro do Corval	6	2008	7,9	393	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	S. Pedro do Corval	3	2007	7,2	343	<Lq	5	<Lq	<Lq	0,44	2,4	nd	nd	10,1	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	S. Pedro do Corval	10	2008	8,1	460	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Sto António do Baldio	3	2007	7	321	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Sto António do Baldio	3	2008	7,9	415	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Sto António do Baldio	10	2008	7,9	453	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Sto António do Baldio	3	2009	7,6	370	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Sto António do Baldio	10	2009	7,3	449	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Telheiro	3	2008	7,7	480	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Telheiro	7	2008	7,6	530	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Telheiro	3	2009	7,7	400	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Telheiro	6	2009	7,6	500	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Reguengos de Monsaraz	Telheiro	2	2007	7,5	330	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Landeira	6	2007	5,9	158	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Landeira	6	2008	5,8	157	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Landeira	6	2009	6,3	160	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Landeira	5	2007	6,8	166	<Lq	3	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Vendas Novas	Landeira	9	2007	6,7	163	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd

Concelho	Localidade/Freguesia	mês	Ano	pH	Condutividade	Amónio	Nitrato	Nitrito	Ferro	Turvação	Oxidabilidade	Potássio	Sódio	Manganês	Cádmio	Crómio	Cobre	Níquel	Chumbo	Alumínio
Vendas Novas	Landeira	1	2008	6,7	166	<Lq	2,9	<Lq	<Lq	<Lq	<Lq	2,9	21	<Lq	<Lq	<Lq	0,006	<Lq	<Lq	nd
Vendas Novas	Landeira	5	2008	7,1	166	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Landeira	2	2009	7	167	<Lq	2,8	<Lq	<Lq	<Lq	<Lq	3,1	21	<Lq	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Vendas Novas	Landeira	5	2009	6,9	165	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Piçarras	5	2007	6,9	227	<Lq	1,7	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Vendas Novas	Piçarras	9	2007	6,8	232	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Piçarras	6	2008	6,7	234	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Piçarras	10	2008	6,8	239	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Piçarras	6	2009	6,7	246	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Piçarras	11	2009	6,4	199	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Vendas Novas	2	2007	7	304	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Vendas Novas	3	2007	7,2	304	<Lq	4,7	<Lq	80	0,5	<Lq	nd	nd	7,9	nd	nd	nd	nd	nd	nd
Vendas Novas	Vendas Novas	4	2007	7,6	286	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Vendas Novas	5	2007	7,3	288	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Vendas Novas	7	2007	7,1	303	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Vendas Novas	8	2007	7,2	290	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Vendas Novas	10	2007	7,2	313	<Lq	5	<Lq	<Lq	0,37	<Lq	2,6	27	6,9	<Lq	<Lq	0,012	<Lq	<Lq	nd
Vendas Novas	Vendas Novas	11	2007	7,2	316	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Vendas Novas	1	2008	7,2	290	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Vendas Novas	5	2008	7,5	288	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Vendas Novas	9	2008	7,4	286	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Vendas Novas	2	2009	7,4	285	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Vendas Novas	5	2009	6,7	289	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Vendas Novas	10	2009	7,2	295	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Vendas Novas	7	2007	5,9	162	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Vendas Novas	10	2008	6,3	179	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vendas Novas	Vendas Novas	11	2009	6	173	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Aguiar	6	2008	7,8	560	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Aguiar	6	2009	7,4	600	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Aguiar	9	2009	7,7	690	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Aguiar	7	2008	7,8	570	<Lq	10,5	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Aguiar	3	2007	7,7	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Aguiar	6	2007	7,8	670	<Lq	14	<Lq	60	<Lq	<Lq	nd	nd	7,3	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Aguiar	1	2007	7,8	500	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Aguiar	5	2007	7,8	550	<Lq	9,5	<Lq	<Lq	0,34	0,99	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Aguiar	10	2007	7,7	640	<Lq	13	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Alcaçovas	5	2009	7,3	570	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Alcaçovas	2	2009	7,5	510	<Lq	9,4	<Lq	<Lq	<Lq	1	3,1	28	<Lq	<Lq	<Lq	0,009	<Lq	<Lq	nd
Viana do Alentejo	Alcaçovas	1	2007	7,6	540	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Alcaçovas	3	2007	7,5	520	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Alcaçovas	5	2007	7,7	500	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Alcaçovas	10	2007	7,5	640	<Lq	13	0,04	<Lq	1,4	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Alcaçovas	9	2009	7,7	650	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Alcaçovas	6	2008	7,6	480	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Alcaçovas	6	2007	7,5	570	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Alcaçovas	3	2008	7,5	590	<Lq	11,5	<Lq	<Lq	<Lq	0,92	2,3	25	<Lq	<Lq	<Lq	<Lq	9	<Lq	nd
Viana do Alentejo	Alcaçovas	9	2008	7,4	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Viana do Alentejo	2	2009	7,2	378	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Viana do Alentejo	5	2009	7	580	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Viana do Alentejo	7	2007	7,3	560	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Viana do Alentejo	3	2008	7,3	560	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Viana do Alentejo	2	2007	7,4	500	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Viana do Alentejo	8	2007	7,2	700	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Viana do Alentejo	11	2007	7,3	620	<Lq	14	<Lq	<Lq	0,34	<Lq	1,8	16	9	<Lq	<Lq	0,009	<Lq	<Lq	nd
Viana do Alentejo	Viana do Alentejo	6	2008	7,2	660	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Viana do Alentejo	9	2009	7,3	690	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Viana do Alentejo	1	2007	7,2	890	<Lq	92	<Lq	<Lq	<Lq	<Lq	nd	nd	6,5	nd	nd	nd	nd	nd	nd

Concelho	Localidade/Freguesia	mês	Ano	pH	Condutividade	Amónio	Nitrato	Nitrito	Ferro	Turvação	Oxidabilidade	Potássio	Sódio	Manganês	Cádmio	Crómio	Cobre	Níquel	Chumbo	Alumínio
Viana do Alentejo	Viana do Alentejo	4	2007	7,3	840	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Viana do Alentejo	10	2007	7,2	870	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Viana do Alentejo	1	2007	7,3	850	<Lq	88	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Viana do Alentejo	4	2007	7,3	820	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Viana do Alentejo	7	2007	7,2	870	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Viana do Alentejo	10	2007	7,3	850	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Viana do Alentejo	9	2008	7,3	620	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Viana do Alentejo	Viana do Alentejo	5	2007	7,3	530	<Lq	10,9	<Lq	<Lq	<Lq	1,08	nd	nd	2	nd	nd	nd	nd	nd	nd
Vila Viçosa	Bencatel	12	2007	7,4	580	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Bencatel	1	2008	7,5	580	<Lq	27	<Lq	<Lq	<Lq	<Lq	1,8	11	<Lq	<Lq	<Lq	0,007	<Lq	<Lq	nd
Vila Viçosa	Bencatel	6	2008	7,4	550	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Bencatel	9	2008	7,4	560	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Bencatel	2	2009	7,2	560	<Lq	28	<Lq	<Lq	<Lq	<Lq	2,7	13	<Lq	<Lq	<Lq	0,007	<Lq	<Lq	nd
Vila Viçosa	Bencatel	6	2009	7,5	540	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Bencatel	10	2009	7,3	560	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Pardais	3	2007	7,4	460	<Lq	15	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Vila Viçosa	Pardais	11	2007	7,6	450	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Pardais	10	2009	7,4	412	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Pardais	4	2008	7,6	439	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Pardais	10	2008	7,4	274	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Pardais	4	2009	7,2	270	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	São Romão	12	2007	7,7	660	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	São Romão	4	2008	7,9	660	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	São Romão	10	2008	7,6	620	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	São Romão	4	2009	7,8	660	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	São Romão	10	2009	7,4	630	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	São Romão	1	2007	7,6	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Vila Viçosa	11	2007	7,4	510	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Vila Viçosa	6	2008	7,4	520	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Vila Viçosa	6	2009	7,4	520	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Vila Viçosa	1	2007	7,6	640	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Vila Viçosa	5	2007	7,7	650	<Lq	24	<Lq	<Lq	<Lq	<Lq	1,7	15	<Lq	<Lq	<Lq	<Lq	<Lq	<Lq	nd
Vila Viçosa	Vila Viçosa	3	2009	7,6	620	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Vila Viçosa	3	2007	7,5	464	<Lq	15	<Lq	<Lq	<Lq	<Lq	nd	nd	<Lq	nd	nd	nd	nd	nd	nd
Vila Viçosa	Vila Viçosa	11	2009	7,4	530	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Vila Viçosa	10	2008	7,5	440	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
Vila Viçosa	Vila Viçosa	3	2008	7,6	480	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd